

GenCore version 5.1.6
Cpypyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 6, 2003, 10:02:23 ; Search time 4075.7 Seconds
(without alignments)
10068.203 Million cell updates/sec

Title: US-(9-775-879-20_COPY_143_1552
Perfect score: 1410
Sequence: 1 atgcgcggaataatgctta.....tgcacatggaacatctg 1410

Scoring table: IDENTITY_NUC
Gapcp 10.0, Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb_ba.*
- 2: gb_htg.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vl.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_om.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_sts.*
- 28: em_un.*
- 29: em_vl.*
- 30: em_htg_hum.*
- 31: em_htg_inv.*
- 32: em_htg_other.*
- 33: em_htg_mus.*
- 34: em_htg_pln.*
- 35: em_htg_rtd.*
- 36: em_htg_mam.*
- 37: em_htg_vrt.*
- 38: em_sy.*
- 39: em_htgo_hum.*
- 40: em_htgo_mus.*
- 41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1410	100.0	2294	8	AF105034	AF105034 Arabidops
2	1397.4	99.1	3004	8	AF069468	AF069468 Arabidops
3	1397.4	99.1	103904	8	AFAC021640	AC021640 Arabidops
4	307.8	21.8	1175	8	AFNASC5D	X90454 A.thaliana
5	307	21.8	1163	8	AFNASC5D	AY088527 Arabidops
6	202.8	14.4	179428	8	AF003214	AP003214 Oryza sat
7	186.2	13.2	1155	8	AF081794	AF081794 Nicotiana
8	178.2	12.6	1148	8	AF099969	AF099969 Nicotiana
9	164.2	11.6	1381	6	AX155063	AX155063 Sequence
10	92.6	6.6	293	11	G71681	G71681 A85715834FM
11	65.8	4.7	310	11	G71612	G71612 A62988234FM
12	57.6	4.1	840	8	CNS0180K	AL110675 Botrytis
13	57.4	4.1	696	8	CNS018D5	AL114081 Botrytis
14	56.6	4.0	127709	2	CNS057C1	AL352983 Homo sapi
15	55.8	4.0	52359	2	AC010772	AC010772 Homo sapi
16	55.6	3.9	7218	6	I66494	I66494 Sequence 14
17	55.6	3.9	166050	2	AC013817	AC013817 Homo sapi
18	55.6	3.9	177008	2	AC011175	AC011175 Homo sapi
19	55.2	3.9	221133	2	AC122896	AC122896 Mus muscu
20	54.8	3.9	6485	6	AX251313	AX251313 Sequence
21	54.8	3.9	6485	6	AX346710	AX346710 Sequence
22	54.4	3.9	224448	2	PFMAL4P4	AL035477 Plasmodiu
23	54.2	3.8	6015	6	AX345578	AX345578 Sequence
24	54	3.8	7008	6	AX251285	AX251285 Sequence
25	54	3.8	114897	2	AP003624	AP003624 Oryza sat
26	53.8	3.8	210105	2	AC125948	AC125948 Rattus no
27	53.6	3.8	107739	2	AC116979	AC116979 Dictyoste
28	53.6	3.8	170568	2	AC107605	AC107605 Rattus no
29	53.4	3.8	83440	2	AC024285	AC024285 Homo sapi
30	53.4	3.8	137216	2	AC055869	AC055869 Homo sapi
31	53.4	3.8	273807	2	AC025421	AC025421 Homo sapi
32	53.2	3.8	131271	2	AC015927	AC015927 Homo sapi
33	53	3.8	152878	3	CXY18D10A	AL034393 Caenorhab
34	52.8	3.7	64612	2	AC079333	AC079333 Homo sapi
35	52.8	3.7	147119	2	AC036177	AC036177 Homo sapi
36	52.8	3.7	147962	2	AC130632	AC130632 Rattus no
37	52.8	3.7	156550	2	AC015830	AC015830 Homo sapi
38	52.6	3.7	3364	14	TSP418778	AJ418778 Tomato sp
39	52.6	3.7	143585	2	AC013349	AC013349 Homo sapi
40	52.2	3.7	14867	3	AE001398	AE001398 Plasmodiu
41	52	3.7	9539	6	AX277889	AX277889 Sequence
42	52	3.7	9539	6	AX323566	AX323566 Sequence
43	52	3.7	163547	2	AC111523	AC111523 Rattus no
44	52	3.7	265985	2	AC087226	AC087226 Mus muscu
45	51.8	3.7	1141	6	AX083744	AX083744 Sequence

ALIGNMENTS

RESULT 1
AF105034
LOCUS
DEFINITION Arabidopsis thaliana delta7 sterol C-5 desaturase (STL) gene,
complete cds.
ACCESSION AF105034
VERSION AF105034.1 GI:5031218
KEYWORDS
SOURCE Arabidopsis thaliana.
ORGANISM Arabidopsis thaliana.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 2294)
AUTHORS Choe,S., Noguchi,T., Fujioka,S., Takatsuto,S., Tissier,C.P.,

Gregory,B.D., Ross,A.S., Tanaka,A., Yoshida,S., Tax,F.E. and Feldmann,K.A.	
The Arabidopsis dwf7/stel mutant is defective in the delta7 sterol C-5 desaturation step leading to brassinosteroid biosynthesis	
JOURNAL	Plant Cell 11 (2), 207-221 (1999)
MEDLINE	99128169
PUBMED	9927039
REFERENCE	2 (bases 1 to 2294)
AUTHORS	Choe,S., Tanaka,A., Gregory,B.D. and Feldmann,K.A.
TITLE	Direct Submission
JOURNAL	Submitted (06-NOV-1998) Plant Sciences, The University of Arizona, Forbes Hall 303, Tucson, AZ 85721, USA
FEATURES	Location/Qualifiers
source	1..2294
	/organism="Arabidopsis thaliana"
	/cultivar="Ws-2 (Wassilewskija-2)"
	/db_xref="taxon:3702"
	/chromosome="III"
	/map="2.2"
gene	43..1889
	/gene="STEL"
	/note="DWARF7"
mRNA	join(43..368,736..1041,1396..1889)
	/gene="STEL"
	/product="delta7 sterol C-5 desaturase"
	join(143..368,736..1041,1396..1709)
CDS	/gene="STEL"
	/note="steroid/brassinosteroid biosynthetic enzyme"
	/codon_start=1
	/product="delta7 sterol C-5 desaturase"
	/protein_id="AAD38120.1"
	/db_xref="GI:5031219"
	/translation="MAADNAYLMQFVDETSFYNRIVLSHLLPANLMEPLPHFLTQTLWR NYLAGTLIFYISGLFMCFFYIYLKINVYLPKDAIPTIKAMRLQMFAMKAMPWYTLLP TVSEMIERQWTKCFASIDBFGLYFVYIAIYLVFVEFGIYWMHRELHDLKPLYKL HATHTYINKONTLSPFAGLPHVDGLQAPHVIALFIVPIHFTTHIGLFPWEATV ANTHDCIHGNIWPMVGAGYHTIHHTYIKHNYGHTIWMDMVFGSLRDLPLEEDNDKDS FKKAE"
variation	322
	/gene="STEL"
	/note="dwf7-2 (stel-3) allele"
	/replace="a"
variation	1552
	/gene="STEL"
	/note="dwf7-1 (stel-2) allele"
	/replace="a"
BASE COUNT	585 a 469 c 417 g 823 t
ORIGIN	
Query Match	100.0%; Score 1410; DB 8; Length 2294;
Best Local Similarity	100.0%; Pred. No. 4.7e-296;
Matches 1410; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
Qy	1 ATGCGGCGGATATGCTTATCTGATGCAGTTGTGTTGACGAAACCTCTTTTACACCGA 60
Db	143 ATGCGGCGGATATGCTTATCTGATGCAGTTGTGACGAAACCTCTTTTACACCGA 202
Qy	61 ATGCTTCTGAGTCATCTTTTCCGGCGGAATCTATGGGAACCCCTACCTCAATTTCTCCAG 120
Db	203 ATGCTTCTGAGTCATCTTTTCCGGCGGAATCTATGGGAACCCCTACCTCAATTTCTCCAG 262
Qy	121 ACATGGCTCCGAATACCTTCGCGGAAACCCCTACTACTCTCCGGTTCCCTCTGG 180
Db	263 ACATGGCTCCGAATACCTTCGCGGAAACCCCTACTACTCTCCGGTTCCCTCTGG 322
Qy	181 TGCCTTCTACATCTATTACCTTAAATCAACGTTTACCTTCCCAAGGCTCTCGACTTTTCAC 240
Db	323 TGCCTTCTACATCTATTACCTTAAATCAACGTTTACCTTCCCAAGGCTCTCGACTTTTCAC 382
Qy	241 TTTTGTATTCATATGCTTAAATCGGCTTCTATGTTATCGATTTTCAATTTAAGGAAGA 300
Db	383 TTTTGTATTCATATGCTTAAATCGGCTTCTATGTTATCGATTTTCAATTTAAGGAAGA 442

Db	1523	CATCCATGACTGCATCCATGGCAACATCTG	1552	Db	1542	AGGGTTTCTTCGTGTACTGTACAGTAATTTGGATTTGATGTGGATAGTTCATGTTTGCAT	1601
RESULT 2	AF069468	3004 bp	DNA	linear	PLN 19-MAY-1999		
LOCUS	Arabidopsis thaliana sterol-C5-desaturase gene, complete cds.						
DEFINITION	AF069468						
ACCESSION	AF069468.1	GI:4234767					
VERSION							
KEYWORDS	Arabidopsis thaliana.						
SOURCE	Arabidopsis thaliana.						
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliopsida; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.						
REFERENCE	1 (bases 1 to 3004)						
AUTHORS	Husselstein, T., Schaller, H., Gachotte, D. and Benveniste, P.						
TITLE	Delta7-sterol-C5-desaturase: molecular characterization and functional expression of wild-type and mutant alleles						
JOURNAL	Plant Mol. Biol. 39 (5), 891-906 (1999)						
MEDLINE	99273992						
PUBMED	10344195						
REFERENCE	2 (bases 1 to 3004)						
AUTHORS	Husselstein, T., Schaller, H., Gachotte, D. and Benveniste, P.						
TITLE	Direct Submission						
JOURNAL	Submitted (30-MAY-1998) IBMP, 28, rue Goethe, Strasbourg 67083, France						
FEATURES	Location/Qualifiers						
Source	1..3004						
	/organism="Arabidopsis thaliana"						
	/cultivar="columbia"						
	/db_xref="taxon:3702"						
	/tissue_type="leaves"						
	join(1242..1467,1836..2141,2496..2809)						
	/codon_start=1						
	/product="sterol-C5-desaturase"						
	/protein_id="AADI2944.1"						
	/db_xref="GI:4234768"						
	/translation="MAADNAYLMQFVDETSFYNRIVLSHLPLANLPELPHFLTWLRLNYLALTYLISGLWCFYIYLLKINVLKPDALPTIKAMRLONFVAMKAMPWYLLLP TVSSMIRGHTKCFASIGEGWILYFYVYIALYLVFERGIYWHRELHDIKPLYKL HATHIYKNQTLSPFAGLAPHPDGIQVPHVIALFIVPIHFTTHIGLLFPEALWT ANHIDCHGNLWPNVWAGYHYIHHTYKHYNGHYTIWMDNMFGLSLRDLLEEDNDKDS FKAE"						
CDS	865 .. 577 c	490 g	1072 t				
BASE COUNT							
ORIGIN							
Query Match	99.1%;	Score 1397.4;	DB 8;	Length 3004;			
Best Local Similarity	99.9%;	Pred. No. 2.4e-293;					
Matches 1409;	Conservative 0;	Mismatches 1;	Indels 1;	Gaps 1;			
QY	1	ATGGCG(CGGGATAAGCTTATCTGATGAGTTCTGTGACGAAACCTCTTTTACACCGA 60					
Db	1242	ATGGCG(CGGGATAAGCTTATCTGATGAGTTCTGTGACGAAACCTCTTTTACACCGA 1301					
QY	61	ATCGTT(TGAGTCATCTTTTCCGCGCGGAATCTATGGGAACCTTACCTCATTTTCTCCAG 120					
Db	1302	ATCGTT(TGAGTCATCTTTTCCGCGCGGAATCTATGGGAACCTTACCTCATTTTCTCCAG 1361					
QY	121	ACATGG(TCCGAAATACCTCGCGGAACCTCTATCTATCTCATCTCCGTTTCTCTCG 180					
Db	1362	ACATGG(TCCGAAATACCTCGCGGAACCTCTATCTATCTCATCTCCGTTTCTCTCG 1421					
QY	181	TGCTTC(ACATCTATTACCTTAAATCAACCTTTACCTTCCCAAGGCTCTCGACTTTCAC 240					
Db	1422	TGCTTC(ACATCTATTACCTTAAATCAACCTTTACCTTCCCAAGGCTCTCGACTTTCAC 1481					
QY	241	TTTTGT(ATTCACTATTGCTTAACTGCTTCTATGTTATCTGATTTTCAATTTT-AAGGAAG 299					
Db	1482	TTTTGT(ATTCACTATTGCTTAACTGCTTCTATGTTATCTGATTTTCAATTTTAAAGGAAG 1541					
QY	300	AGGGTT(CTTCGTGACTGTACAGTAATTTGGATTTTGTGATGTTGATGTTTGCATCTGATGTTTGCAT 359					

Db 2622 ACATCCATGACTGCATCCATCCATGCGCAACATCTG 2652

RESULT 3

ATAC021640 103904 bp DNA linear PLN 24-JAN-2001

LOCUS Arabidopsis thaliana chromosome III BAC F16B3 genomic sequence, complete sequence.

DEFINITION

AC021640

AC021640.7 GI:12408747

HTG

Arabidopsis thaliana

SOURCE

ORGANISM

Arabidopsis thaliana

Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi.

1 (bases 1 to 103904)

Lin,X., Kaul,S., Town,C.D., Benito,M.-I., Creasy,T.H., Haas,B., Wu,D., Ronning,C.M., Koo,H., Fujii,C.Y., Utterback,T.R., Barnstead,M.E., Bowman,C.L., White,O., Nierman,W.C. and Fraser,C.M.

Arabidopsis thaliana chromosome III BAC F16B3 genomic sequence

Unpublished

2 (bases 1 to 103904)

Lin,X. and Kaul,S.

Direct Submission

REFERENCE

AUTHORS

TITLE

JOURNAL

3 (bases 1 to 103904)

Medical Center Dr, Rockville, MD 20850, USA, xlinetigr.org

Lin,X.

Direct Submission

Submitted (24-JAN-2001) The Institute for Genomic Research, 9712 Medical Center Dr., Rockville, MD 20850, USA

On Jan 24, 2001 this sequence version replaced gi:12280770.

Address all correspondence to:atetigr.org

BAC clone F16B3 is from Arabidopsis chromosome III and is near the molecular marker mi74.

The orientation of the sequence is from SP6 to T7 end of the BAC clone.

Genes were identified by a combination of three methods: Gene prediction programs including GRAIL (available by anonymous ftp from arthur.epm.ornl.gov), Genefinder (Phil Green, University of Washington), Genscan (Chris Burge, <http://gnomic.stanford.edu/~chris/GENSCANW.html>), and NetPlantGene (<http://www.cbs.dtu.dk/netgene/cbsnetgene.html>), searches of the complete sequence against a peptide database and the Arabidopsis EST database at TIGR (<http://www.tigr.org/tdb/at.html>).

Annotated genes are named to indicate the level of evidence for their annotation. Genes with similarity to other proteins are named after the database hits. Genes without significant peptide similarity but with EST similarity are named as 'unknown' proteins. Genes without protein or EST similarity, that are predicted by more than two gene prediction programs over most of their length are annotated as 'hypothetical' proteins. Genes encoding tRNAs are predicted as tRNAst-SE (Sean Eddy, <http://genome.wustl.edu/eddy/tRNAst-SE/>). Simple repeats are identified by repeatmasker (Arian Smit, <http://ftp.genome.washington.edu/RM/RepeatMasker.html>).

FEATURES

source

1..103904

/organism="Arabidopsis thaliana"

/cultivar="Columbia"

/db_xref="taxon:3702"

/chromosome="III"

/map="mi74"

/clone="F16B3"

<129..>979

/gene="F16B3.1"

/note="predicted by genefinder"

join(<129..241,586..>979)

/gene="F16B3.1"

join(129..241,586..979)

/gene="F16B3.1"

gene

note="hypothetical protein"

/codon_start=1

/protein_id="AAF32448.1"

/db_xref="GI:6957704"

/translation="MQLSPFCRPMSENREANSSEEMELLHGSNRLSPSEHVRVYSGNSSEDGSPRICRQOFSRIGHAAETYLITLRSFNLLGLYLGDLIIPPTSDGLKPVVYVTGAWIIGYKAWGLQLAERDIIIVCLDYRNFPQGTISDMVSDAAQGSFVGNISAFGDPNRIYLMGOSAGAHISSCALFQAIKESGESISVSQIKAYFGLSGYNLENLVEHFNRLYRIFLSIMEGESPKQFSEVRLKDLNVRKAAALPHILFGNSADYSIPPEASKTFTDALQAEVKALVYKGTHTDILQDPLRGCKDELFDHIVSMHADDSDALRNDVAPRRKLVEFLKLKAGRVSP"

note="hypothetical protein"

/codon_start=1

/protein_id="AAF32445.1"

/db_xref="GI:6957701"

/translation="MEEKREGSPLESENEAGLLCYDDEYVQGGPAPKQLQRYLSEIGELQIGDEDVVISLADLYGAETAGKYGCCWENYEVRLKGLGLYILGRHGVPTWKYAVNTPTSDDESCLCAAEFFQDRDSVTKLLSDMHICDARPVFDVYLPNSFKKSSGPESVTCFSG"

complement(<3079..>4265)

/gene="F16B3.2"

/note="identical to COL2 CB:AAB67879 from [Arabidopsis thaliana]"

complement(join(<3079..3438,3582..>4265))

/gene="F16B3.2"

complement(join(3079..3438,3582..4265))

/gene="F16B3.2"

/codon_start=1

/product="COL2"

/db_xref="GI:6957702"

/translation="MLKESNESGTWARACTCRSACTVYCEADSAIYLCITCTDARVHLAARVASRHERVRCOCESAPAEFLCKADAAISLCTACDAEIHSAANLARRHORVPLPLSANSCTSMAPSETDADNEDDDREVASMLLPNGKNIGNONNGFLGVEYLDLDVYSSSMDNFEDNQHITHORSGFGGVPLQVEESTSHLQSOONFOLGINTGFGSGAHYNNSLKDNLNHSASVSSMDISVVPESASTDITVQHPRTTKETIDLSGPPQVVOOLTPMEREARYLRYREKKTKRKFDKTIYASRKAYAEIRPRIKGRFAKRIETAEAEAEFISTLSMETGYGVPSF"

5228..5255

/rpt_family="(CAA)n"

<5471..>7228

/gene="F16B3.3"

/note="predicted by genefinder"

<5471..>7228

/gene="F16B3.3"

5471..7228

/gene="F16B3.3"

note="hypothetical protein"

/codon_start=1

/protein_id="AAF32447.1"

/db_xref="GI:6957703"

/translation="MVPPLKLDFTQPRAGDSLGFKPGSTIRIGTRVGRNELAIKDA GSTKHLRIVSDSENIILHIGSSNGTILSDTIDSDTVPNLSGDEIKLGVEITSLVNGSDVQAPQEHKLPFRPRNRNRLAASDPDPDPIESVQEKPRKRGSKQENELP KTSRARKNLDIADKDELDVEIEKVARVRPRKNAGSAIAKEEVEVEKKRVGRPNKASSAITEEEVEEKGSRARRKNSIYQKSIKLEVEDTPKAVEISEVKSRRKVTYSKQIENECFLEVEDEKTRTSKTTEIGESFLEMLWLNQARKSRARRKMDPEPSKETRNDAGEVLKNCHEEDKNEAQEGCSGRSDDICQEDDEKCDGSKRVQVETELRKKSTVEGLKCTDKEDGETENSODLERENENEKEAQEGCSERSDKVE RVGGAKRVEQVEIELRKKSTVGEDDLNCTVRDGETENLQETEECHDESCKVEE AGFATLDEKVGQGWNNKVERVEVYLEKMKLREWFDAIEVQLPKOTIETEKMIEPM RSKSMRVHRIAEQKEKGD"

complement(<7759..>10518)

/gene="F16B3.4"

note="predicted by genscan, contains Coesterase domain"

complement(join(<7759..7920,8001..8091,8192..8335,8796..8861,8964..9101,9182..9283,9433..9507,9638..9718,10256..>10518))

/gene="F16B3.4"

complement(join(7759..7920,8001..8091,8192..8335,8796..8861,8964..9101,9182..9283,9433..9507,9638..9718,10256..10518))

/gene="F16B3.4"

note="hypothetical protein"

/codon_start=1

/protein_id="AAF32448.1"

/db_xref="GI:6957704"

/translation="MQLSPFCRPMSENREANSSEEMELLHGSNRLSPSEHVRVYSGNSSEDGSPRICRQOFSRIGHAAETYLITLRSFNLLGLYLGDLIIPPTSDGLKPVVYVTGAWIIGYKAWGLQLAERDIIIVCLDYRNFPQGTISDMVSDAAQGSFVGNISAFGDPNRIYLMGOSAGAHISSCALFQAIKESGESISVSQIKAYFGLSGYNLENLVEHFNRLYRIFLSIMEGESPKQFSEVRLKDLNVRKAAALPHILFGNSADYSIPPEASKTFTDALQAEVKALVYKGTHTDILQDPLRGCKDELFDHIVSMHADDSDALRNDVAPRRKLVEFLKLKAGRVSP"

698	CGTGGTTGGACCAAAATGTTTTGCTAGCATAGACGAATTCGGCTGGGATTCGTATTGTTGTT 757	QY
448	CGTGGTTGGACCAAAATGTTTTGCTAGCATAGGGAATTCGGTTGGATTCTGTATTGTTGTT 507	DB
758	TACATCGCCATCTAICTTCTGTTTTCGTTGAGTTGGTATTATTGATGCACAGAGCCTT 817	QY
508	TACATCGCCATCTAICTTCTGTTTTCGTTGAGTTGGTATTATTGATGCACAGAGCCTT 567	DB
818	CATGACATTAAGCCTCTCTATAGTATCTCCATGCCACCCATCATATCTACAAACAGCAG 877	QY
568	CATGACATTAAGCCTCTCTATAGTATCTCCATGCCACCCATCATATCTACAAACAGCAG 627	DB
878	AATACACTCTCTCCATTTGCCGGTAAAGTGTGTTTCA 912	QY
628	AATACACTCTCTCCATTTGCCGGCTTGCATTTC 662	DB
AY088527	Arabidopsis thaliana clone 7564 mRNA, complete sequence.	
AY088527	1163 bp	
AY088527	1 GI:21407301	
FLI_CDNA		
thale cress.		
Arabidopsis thaliana		
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;		
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi		
1 (bases 1 to 1163)		
Haas,B.J., Volkovskiy,N., Town,C.D., Troukhan,M., Alexandrov,N.,		
Feldmann,K.A., Flavell,R.B., White,O. and Salzberg,S.L.		
Full-length messenger RNA sequences greatly improve genome		
annotation		
Genome Biol. (2002) In press		
2 (bases 1 to 1163)		
Brover,V., Troukhan,M., Alexandrov,N., Lu,Y.-P., Flavell,R. and		
Feldmann,K.		
Full-length cDNA from Arabidopsis thaliana		
Unpublished		
3 (bases 1 to 1163)		
Brover,V., Troukhan,M., Alexandrov,N., Lu,Y.-P., Flavell,R. and		
Feldmann,K.		
Direct Submission		
Submitted (11-MAR-2002) Ceres, Inc, 3007 Malibu Canyon Road,		
Malibu, CA 90265, USA		
This clone sequence is one of 5,000 Ceres full-length cDNAs made		
available to TIGR and Genbank. The following quality assessment of		
this set was done by comparison with known proteins: two percent of		
the clones are estimated to be 5'-truncated; less than one percent		
are 3'-truncated; approximately two percent represent alternative		
splice variants, including unspliced introns and spliced exons; one		
percent may contain premature stop codons; five percent may have		
frame shifts in a coding region. A sequence is considered to be		
5'-truncated if it lacks the translation initiation start (ATG). A		
sequence is considered to be 3'-truncated if it lacks the		
C-terminal end of the encoded protein. Please note that these cDNA		
sequences are derived from the Ws or Laer ecotypes and therefore		
may contain polymorphisms when compared to sequences from Col-0.		
Genet carried out the library production and sequencing of the		
full-length clones. Ceres, Inc. carried out the clustering of the		
5' sequences, selection of clones, and sequence assembly.		
Location/Qualifiers		
1. .1163		
/organism="Arabidopsis thaliana"		
/db_xref="taxon:3702"		
/clone="7564"		
146. .991		
/codon_start=1		
/product="sterol-C5-desaturase"		
/protein_id="AA066060.1"		
/db_xref="GI:21594971"		
/translation="MAADNAYLMQFVDETSFYNRIVLSHLLPANLWEPHPFLQWLRL		
NYLAGLLLIISGFCMCFYIYKQFVDETSFYNRIVLSHLLPANLWEPHPFLQWLRL		
THSEINRGWKTGIFGWLIVYAIYLVFEFGIYMHRELHDIRLKYLL		
HATPHYIKNTLSPPAGLAFIPVDGILQAVPHIALFTVPIHFTTHIGLLEWELWA		
ANHCDCIHGNIPVMGAGVTHIHTTYYKKNHYTIIMDMFGLRDLPLEEDNNKS		
FKAE"		
BASE COUNT 300 a 265 c 227 g 383 t		
ORIGIN		
Query Match. 21.8%; Score 307.8; DB 8; Length 1175;		
Best Local Similarity 94.9%; Pred. No. 1.2e-56;		
Matches 318; Conservative 0; Mismatches 17; Indels 0; Gaps 0;		
QY 578 GTCTGTTTTCCTAGATGCAATTCCTACAAAGGCTATGCGTTGCAATGTTTGTG 637		
DB 328 GTTTACCTTCCCAAGATGCAATTCCTACAAAGGCTATGCGTTTGCATGTTGTG 387		
QY 638 GCAATGAAGCTATGCGATGCTACATCTCTCCAACTGCTCCCGAGAGTATGATTGAA 697		
DB 388 GCAATGAAGCTATGCGATGCTACATCTCTCTCCAACTGCTCCCGAGAGTATGATTGAA 447		
63889	CAAAGGATGACATCTATATATTACTTGTAACTTCCTTAGCTTTTGTGTACAGGCGCTTG 63948	QY
1260	CATTTCCACCCAGTAGACGGGATACTTACCGCTGTACCGCATGTCATAGCGCTGTTTATAG 1319	QY
63949	CAITTCACCCAGTAGACGGGATACTTACCGCTGTACCGCATGTCATAGCGCTGTTTATAG 64008	DB
1320	TGCCAATTCATTTCAACAATCATATAGTCTTTTGTTCATGGAAGCGATATGACGCGCGA 1379	QY
64009	TGCCAATTCATTTCAACAATCATATAGTCTTTTGTTCATGGAAGCGATATGACGCGCGA 64068	DB
1380	ACATCCATGACTGCATCCATGGCAACATCTG 1410	QY
64069	ACATCCATGACTGCATCCATGGCAACATCTG 64099	DB
RESULT 4		
ATRNASC5D		
LOCUS		
DEFINITION		
ACCESSION		
VERSION		
KEYWORDS		
SOURCE		
ORGANISM		
Arabidopsis thaliana		
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;		
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi		
1 (bases 1 to 1175)		
Gachotte,D., Husselstein,T., Bard,M., Lacroute,F. and Benveniste,P.		
Isolation and characterization of an Arabidopsis thaliana cDNA		
encoding a delta 7-sterol-C-5-desaturase by functional		
complementation of a defective yeast mutant		
Plant J. 9 (3), 391-398 (1996)		
97077348		
8919315		
2 (bases 1 to 1175)		
Gachotte,D.		
Direct Submission		
Submitted (12-JUL-1995) D. Gachotte, Institut de Biologie		
Moleculaire Plantes, Enzymologie Moleculaire et Cellulaire, FRANCE		
Institut de botanique 28, rue Goethe, F- 67083 Strasbourg, FRANCE		
Location/Qualifiers		
1. .1175		
/organism="Arabidopsis thaliana"		
/variety="Landsberg erecta"		
/db_xref="taxon:3702"		
118. .963		
/codon_start=1		
/product="sterol-C5-desaturase"		
/protein_id="CAA62079.1"		
/db_xref="GI:1061038"		
/db_xref="SPTREMBL:Q39208"		
/translation="MAADNAYLMQFVDETSFYNRIVLSHLLPANLWEPHPFLQWLRL		

NLAGTLLYFISGLMFCFYIYKINVLKPKDAIPTIKAMRLQMFVAMKAMPWYTLIP
TVSSMIKQKXGKFSAGIGFWMILYFYIYLVFVEFGYIMHRELHDKPLKYKL
HATHSIYKQNTLSPFAGLAPHPVDGILQAVPHVIALFIVPHTHTHGLGLFMEAIWT
ANIHCDCIGNLWPMVGAGYHTIHTTYKHNHYGHTIWMWFMGSLRPLLEEDNDKDS
FKKAE"

BASE COUNT 291 a 263 c 223 g 384 t 2 others

Query Match 21.8%; Score 307; DB 8; Length 1163;
Best Local Similarity 94.3%; Pred. No. 1.8e-56;
Matches 316; Conservative 2; Mismatches 17; Indels 0; Gaps 0;

QY 578 GTCGTGTTTGCATGATGCAATCTCTACATAAAGCGTATCGGTTTCCAAATGTTGTG 637
Db 356 GTTTACCTTCCCAAGATGCAATCTCTACATAAAGCGTATCGGTTTCCAAATGTTGTG 415
QY 638 GCAATGAJGGCTATGCCATGTACACTCTCTTCCCACTGCTCCGAGCATGATTCGAA 697
Db 416 GCAATGAJGGCTATGCCATGTACACTCTCTTCCCACTGCTCCGAGCATGATTCGAA 475
QY 698 CGTGGTTCGACCAATGTTTCTAGCATAGCAATTCGGTGTGATTCCTGATTTTGT 757
Db 476 CGTGGTTCGACCAATGTTTCTAGCATAGCAATTCGGTGTGATTCCTGATTTTGT 535
QY 758 TACATCGCATCATATCTGTTTCTGTTGAGTTTGGTATTTATGATGCACAGAGCGTT 817
Db 536 TACATCGCATCATATCTGTTTCTGTTGAGTTTGGTATTTATGATGCACAGAGCGTT 595
QY 818 CATGACAATAAGCTCTCTATAAGTATCTCCATGCCACCATCATATCTACACAAGCAG 877
Db 596 CATGACAATAAGCTCTCTATAAGTATCTCCATGCCACCATCATATCTACACAAGCAG 655
QY 878 AATACACATCTCTCCATTTCCCGGTAAAGTGTTCCTCA 912
Db 656 AATACACATCTCTCCATTTCCCGGTTCGATTTCA 690

RESULT 6
AP003214
LOCUS
DEFINITION
AP003214
ACCESSION
AP003214
VERSION
AP003214.3
KEYWORDS
SOURCE

179428 bp DNA linear PLN 21-MAR-2002
Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 1,
BAC clone:OSJNBa0083M16.

AP003214.3 GI:15528580

Oryza sativa (japonica cultivar-group) (cultivar:Nipponbare) DNA,
clone:OSJNBa0083M16.

Oryza sativa (japonica cultivar-group)
Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.

1
Sasaki, T., Matsumoto, T. and Yamamoto, K.

Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, BAC

clone:OSJNBa0083M16

Published Only in Database (2001)

2 (bases 1 to 179428)

Sasaki, T., Matsumoto, T. and Yamamoto, K.

Direct Submission

Submitter (19-FEB-2001) Takuji Sasaki, National Institute of

Agricultural Sciences, Rice Genome Research Program, Kannondai

2-1-2, Tsukuba, Ibaraki 305-8602, Japan

(E-mail:tsasaki@nias.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/,

Tel:81-298-38-7441, Fax:81-298-38-7468)

On Sep 7, 2001 this sequence version replaced gi:13620983.

Genes were predicted from the integrated results of the following:

GENSCAN1.0, BLASTN2.0, BLASTX2.0 as well as SplicePredictor

(October 1998 version). The genomic sequence was searched against

NCBI Nonredundant protein database, nr

(ftp://ncbi.nlm.nih.gov/blast/db) and the cDNA sequence database at

RGF. Protein homologues of the coding regions were searched against

NCBI Nonredundant protein database with BLASTP2.0. ESTs represent

the identified cDNA sequences using BLASTN 2.0 with the

corresponding DDBJ accession no. and RGP clone ID.

A gene with identity or significant homology to a protein is
classified based on the protein name to indicate the homology level
such as same name, 'putative-', and '-like protein'. A gene without
significant homology to any protein but with EST homology (covering
almost the entire length of partial sequence) is classified as an
'unknown' protein. A gene predicted with a gene prediction program
is classified as a 'hypothetical' protein.

The orientation of the sequence is from -21M13 to M13rev of the BAC
clone. This sequence of OSJNBa0083M16 clone has an overlap with
P504H10 (DDBJ: AF002526) clone at the position 1 to 42,554 and with
P0443D08 (DDBJ: AP003250) at the position 178,805 to 179,428 of 3'.
end. The sequence of this clone starts at the position 100,962 of
P504H10 and ends at the position 624 of P0443D08. Detailed
information on overlap and assembly quality together with
annotation of this entry is available at
http://rgp.dna.affrc.go.jp/Genomeseq.html.

location/Qualifiers

FEATURES

source

1. 179428

/organism="Oryza sativa (japonica cultivar-group)"

/cultivar="Nipponbare"

/db_xref="taxon:39947"

/chromosome="11"

/clone="OSJNBa0083M16"

complement(join(870..1087,1280..1364))

/gene="OSJNBa0083M16.1"

complement(join(870..1087,1280..1364))

/gene="OSJNBa0083M16.1"

/note="contains EST AU092803(C53725)"

/codon_start=1

/product="similar to Zea mays Bowman-Birk proteinase

inhibitors WIPI precursor"

/protein_id="BAB64603.1"

/db_xref="GI:15528581"

/translation="MKSSTLIVLAVQALMGIFAAYAKENAVCESKAIDNPQLK

CCSCNFSGLYTCDDIVKKCDPVCKCAVVKTPVVKMFKCTDTFLMGCGPPCAH"

5689..9051

/gene="OSJNBa0083M16.2"

5689..9051

/gene="OSJNBa0083M16.2"

/codon_start=1

/product="similar to Lycopersicon pimpinellifolium Cf-2

gene"

/protein_id="BAB64604.1"

/db_xref="GI:15528582"

/translation="MYSDFEPALFSLTSLORLDSMNSLGTSSSTTKDAEFDRLTSLT

HLNLSNGLDQIPMGINKLINLVSLDSKRYVNDSDISFNESDEIFITGDSNHL

QESRLMSLVENLSKELYLDHVDNSTVDDMCKTQAQVPRQLVSLDGLCCSLNPIH

HSLLRLHSITVINGQSPGIAVNFPPDFMGFANLTVLRSHNNLEGGWPPDFQKRN

LRILDLSPNNMLLGHLPKVPPTSLERLGTNFSYAKRTSSNFNMKLKELGKLLS

KDLTSLFGLIWSLCHLELNSLGLDGLNSLWIGAHKNTCLILSEFDFSTKPS

ISNFKNLBSLPGCNLTPIMSAIGDLVQLSDMNCNTYSSMPSISGLNLSKL

YINSPGFLGPPAALGNLSKSWFVSCETGPMPTIGNLTQTLQTLAICRSPGP

IPIYSIGQLKRALFIEGCMNSGRIPNSIVNMSKLIYGLPANYLSGKIPARLTLPA

LLFLDLFGNHSFGTQEDAVPSYLSMLQTLNSETLGEFFKFFELTLIALEIDLNN

LAGSVDLFSFKRLKRLDNLNHNLSVMDDEGDNSSSTYLSLKLGLACNITKF

PSILTRLSDMSYLDLSCNKISGNIPKTIWEKSSVHLNLSHMLTSMVASYLLPF

NRHETLDSNNMLQGOIPNLISAEFDYSHNAFESSILPNFTLYLSKTYLSMKN

ISGNIPHSICSSLLVNLNHNFSGPPSCLMEOQTYFRNLNLRGNFEGMLPNTVT

RCATPTIDLNGKEGRUPRAAGNCTYLEVLIDGNKKTADTFPPMSYLSNLRVLIR

SNRLYSIGTYTFEDKSGDFPNLQIIDLANNSTGSLHPQWFKFISKKYNNKTGTI

SHRISIDGTYQDVTISCKGFSMTFERILTLTAIDLSNALEGISPEISVGLSLH

VNLNLSHNASGRIPPOIGITALESLDLSNNWISGEIPELTNLFLTVLNLNNOLE

GKIPESGFATFENSSYEGNACIGDPLPKCASWPPSPNPHVESSESHVDIVLFLV

complement(join(9685..9961,10319..10662))

/gene="OSJNBa0083M16.3"

complement(join(9685..9961,10319..10662))

/gene="OSJNBa0083M16.3"

/note="hypothetical protein"

/codon_start=1

/protein_id="BAB64605.1"

/db_xref="GI:15528583"

/translation="MESPPRSCSDHAAAIISVMTGRRRLRHMAELAAAREPGSPQLDSL
QSMRSHLGCADGDTADACAGIELHGLAWPRTPPCAATRRRCRPPHPCGAAR
SAAARMA7CNDVSVVLPVPPGAAVNPPLAPASAPSPVRRPQRRPPRSVAT
ICARGGPEERGAAYEGGRRAPAAEVGPAPALGAVAAWCGGG"
complement(join(12004,.12746,13413,.13449))
/gene="OSJNbA0083M16.4"
complement(join(12004,.12746,13413,.13449))
/gene="OSJNbA0083M16.4"
/note="hypothetical protein"
/codon_start=1
/protein_id="BAB64606.1"
/db_xref="GI:15528584"
/translation="MTVLGRKSSGLSVAVAPMEELAPELIGPRISFSDGVVATAGA
ATPASPDSLVVWSSNRRLPEPEFDFAAAADVAPADRLFAGKLLPVPLPPVH
KPSPCQOQATSGGQVKTYSQRRPGSWTSPFTRSCSVNSATTAAPRSGSGSFSCPS
FPLMRSGAAGAGGGLGVSGGHHRRPQHKKAGATAAAYYGGSRNGSSGHG
VRVSPVNPSTGTSNVNMLSYLLDCGKNKTKNGRGLNC"
complement(join(16789,.17270,17559,.17837,17983,.18106,
18193,.18498,18750,.18902))
/gene="OSJNbA0083M16.5"
complement(join(16789,.17270,17559,.17837,17983,.18106,
18193,.18498,18750,.18902))
/gene="OSJNbA0083M16.5"
/note="hypothetical protein"
/codon_start=1
/protein_id="BAB64607.1"
/db_xref="GI:15528585"
/translation="MGRLLLVSLDRRRRVLPSHCQVDPDVVCGWRNAQRTGQH
DYHGFEPKLLPLLSPPRPVPLHPIARAKNSILMEGKHAVDLSSEERBG
GAPGSGSDEVDGAVLGGAESTEPGSSVDDHGVAAEAKNDLDEPLEPE
GVTAQGGDTPADSGENELHEVPKPMATRFAPRATDANATPVYRGMLSRSRK
NAGPTFVGASADPTPALSTSSSELVSSPENASVADPDHAGSPGRNEODL
EAAKQSEARFADAAAAAASREHPSRSKQPRPERFIPERGEAAARAKACRS
GALDRFTSOLNPPSGPSTEWERTAADVVGGOGEPTSDOPSCSIASLSDGPP
EPLPDREIRIVSLAVGLVSLAVSMALVLEYIGSESPSPSPDNQVIEKQV"
join(21657,.22226,22293,.22395,23106,.23221,24403,.24457,
25399,.25589)
/gene="OSJNbA0083M16.6"
join(21657,.22226,22293,.22395,23106,.23221,24403,.24457,
25399,.25589)
/gene="OSJNbA0083M16.6"
/note="hypothetical protein"
/codon_start=1
/protein_id="BAB64608.1"
/db_xref="GI:15528586"
/translation="WISAEHPHPPPPDVADGAAVEAAASARHOAAVLSPFSLK
TWGSHVLCAHVNAGDAIAAARRSPEOLDGKRLLLHLREVDSDDVVADEAVAA
VAAPAPAAASPKLRSRRRNPAPASSPPPERRAARAALDRAREFVTLTSE
ETIEDVYATGARPRRRPRPVOKLDGFRNFKIHVFCYFPFHGCPSPSP
TGCFTIVPGLHSEELRMVDYLTGFGRGPGQVCSASRTMKGIHNVNEHHSF
LSPWNYLEPEVELDALRCQLSCSDCAPRPSRSSRSIYYNQHLQAMFELHQD
RLGGTGICKD"
complement(join(30012,.30245,31517,.31582,31671,.31817))
/gene="OSJNbA0083M16.7"
complement(join(30012,.30245,31517,.31582,31671,.31817))
/gene="OSJNbA0083M16.7"
/note="contains ESTs AU075676(E10683), AU075677(E10683)"
/codon_start=1
/product="similar to Mus musculus SURF-5 gene"
/protein_id="BAB64609.1"
/db_xref="GI:15528587"
/translation="MSKAGAAAGPTAAAAAAYVQKQKTLQAKADVSSLVNFAAF
INIRVNDPVRNTQAFQMDRGRSLIHSADSLKLVLSELKRTAIFSLGSLTENVD
RRIELSQAEQTERLMERIGEAAGSLKELEAHYSSVVRTPPDE"
complement(42717,.43520)
/gene="OSJNbA0083M16.8"
complement(42717,.43520)
/gene="OSJNbA0083M16.8"
/note="hypothetical protein"
/codon_start=1
/protein_id="BAB64610.1"
/db_xref="GI:15528588"
/translation="MSWSPRLKLFQFDVSEETABHEQYSAEPESAPVCGNGGGG

GGSDSSSSSTTTTTATAGDGRREYECQCYREFANSQALGQHNAHKKERQQLKRA
QLQAAAAAGRVAGGAAALYPRANPWSAFAPTPHLLGGEAGPTSWVYFSRAAVAL
QAGQOQFHVSHGCVFARGAGPAVAASPAVFSVTPATSAVAASAPVYTTDEH
HGRRIHASSAASLARYPGAMVVEPVPVPTDCLGLDLQLSLAPAGI"
complement(join(57300,.57361,57501,.57585,57688,.57936))
/gene="OSJNbA0083M16.9"
complement(join(57300,.57361,57501,.57585,57688,.57936))
/gene="OSJNbA0083M16.9"
/note="hypothetical protein"
/codon_start=1
/protein_id="BAB64611.1"
/db_xref="GI:15528589"
/translation="MSLATVAPASRLSLPTSSPEPCRRRLLLSAAAPLRRRLRR
RIAVASVDPDAARAEYPPWLIAGLGNPGSKYHGRHNVGFEMVDRIARDEGTMNT
IQKSLIGIAETJVVEAIDLKTKTLTEV"
complement(59958,.63558)
/gene="OSJNbA0083M16.10"
complement(59958,.63558)
/gene="OSJNbA0083M16.10"
/note="contains EST C26390(C12241)"
Query Match 14.4%; Score 202.8; DB 8; Length 179428;
Best Local Similarity 55.3%; Pred. No. 3.3e-34;
Matches 466; Conservative 0; Mismatches 357; Indels 19; Gaps 3;
QY 585 TTGCTTCAGATGCAATTCCTACAATAAAGGCTATGCGTTTGGCAATGTTTGGCAATGA 644
DB 131331 TTGATGCAGATCTTATACCTACAATCGAAGCTATGAAGAAGCAATAATTGTTGCATCAA 131390
QY 645 AGCTATGCGATGGTACACTCTCTTCCAACTCTCTCCAGAGATGATGTTGAACGTTG 704
DB 131391 AGCTATGCGTCTCTATGTTGCCCTTCCCAACCTTATCTGAGTACATGTTGAGAATGGAT 131450
QY 705 GGACCAAAAGTTTGTGCTAGCATAGACGAATTCGGCTGGATTCGTATTTGTTTACATCG 764
DB 131451 GGACAGAGTTTATGTTATATACGTGAAGTGGTTGGCAATGTCACCTGTTTATCTGG 131510
QY 765 CCATCTATCTCTGTTTTCGTTGAGTTTGGTATTTATTTGATGTCACAGAGCTTCATGACA 824
DB 131511 CTTTATATCTTATCTTTGTTGAGTTTGGAAATTTACTGGATGTCACAGAGTTCATGACA 131570
QY 825 TTAAGCCTCTATAGATATCTCCATGCCACCATCATATCATATCATACACAGCAATACAC 884
DB 131571 TAAAGCCATTGTACAAGTACCTGCACACATACATACACCAATATTTACAACAAGGAGATACC 131630
QY 885 TCTCTCCATTTCGCGTAAAGTGT-----TTTCAGTTTGTCTTCTTTAGTTCT 932
DB 131631 TATCACCATTTCGAGTAAAGTTTATTTGAGCGCGTTTAAATATGCAATTCATCTTTTA 131690
QY 933 TGTAAAGATTTGTTAGCTTTAGTTTCTTACCAAGAAAGACTTTGTCAGCAGCTGCTTGT 992
DB 131691 TGTTCGCTGTGCATAGATATATATGAGAACAGATTCATATTCGGAATGCTTATTTT 131750
QY 993 ACTCCAAATACATTTTGGATTCCTTATCCATAAAGTAAACCAGAGCTAGATATATAT 1052
DB 131751 ACCATGCTTCCAGCTTTTGTAGTATAATCTGTGTATAGTTATATATATTTTCCCTGCTAA 131810
QY 1053 AAATGTCAGCTGATTTACTTTCATATATGTCAGAGAGACTTCTGAC----TTAACACAGAGT 1108
DB 131811 TTAATGTCGTTAAACACTAGGATACCTTATTTAAGGCTCCTAAACCTGTTACCGTAGG 131870
QY 1109 TTAGATCTTTGTTGTTCTCTCTTCGCTCGGACTGATTTGGAATGACGAGAACTTCTTTT 1168
DB 131871 CCATACCAATAAATAAAAAAATGCCATGGCGCTCCCAATGTACATTTTTTAAA---CAATA 131927
QY 1169 ATCTACTTCCCTGGAGTGTATCTTGGTTAATCCAAGATGTGACATCTATATATCTTGT 1228
DB 131928 CAATTTCTTTAGTTAACTCCCTTTGTTGCTGAGATTTGTTATTTCCACCAAGCTTCT 131987
QY 1229 AACTTCTTTAGCTTTTGTGTTTACAGGGCTTGCAATTTCCACCGTAGCAGGATCTTCTAG 1288
DB 131988 GATTTATGCTTGCATGCTTTCCAGGACTAGCATTCATCCACTGATGGATTTTGCAT 132047

QY	1289	GCYGTACCCGATGTGATAGCGCTGTTTATAGTGCACATTCATTTCAACAATATAGT	1348
Db	132048	GCATACCCGATGTGTTTGGCGCTACTCTTATCCCAACACACTTCAGGACACACATTCCT	132107
QY	1349	CTTTGTTTCATGAAGCGATATGACGCGGCAACATCCATGACTGCATCCATGCGCAACATC	1408
Db	132108	CTTTGTTTCATGAAGCGCGGTGGACAACTAACATCCATGACTGCATTCACGCGCAAGTT	132167
QY	1409	TG 1410	
Db	132168	TG 1321f9	
RESULT 7			
LOCUS	AF081794		
DEFINITION	Nicotiana tabacum sterol-C5(6)-desaturase mRNA, complete cds.		
ACCESSION	AF081794		
VERSION	AF081794...	GI:4140397	
KEYWORDS			
SOURCE	Nicotiana tabacum.		
ORGANISM	Nicotiana tabacum.		
REFERENCE			
AUTHORS	Husselstein, T., Schaller, H., Gachotte, D., and Benveniste, P.		
TITLE	Delta7-sterol-C5-desaturase: molecular characterization and functional expression of wild-type and mutant alleles		
JOURNAL	Plant Mol. Biol. 39 (5), 891-906 (1999)		
MEDLINE	99273992		
PUBMED	10344195		
REFERENCE			
AUTHORS	Benveniste, P.		
TITLE	Direct Submission		
JOURNAL	Submitted (30-JUL-1998) Plant Molecular Biology Institute, Centre National de la Recherche Scientifique, 28 rue Goethe, Strasbourg 67084, France		
FEATURES			
source	Location/Qualifiers		
	1..1155		
	/organism="Nicotiana tabacum"		
	/cultivar="xanthi SH6"		
	/db_xref="taxon:4097"		
	/tissue_type="calli derived from leaf protoplasts"		
	103..918		
	/codon_start=1		
	/product="sterol-C5(6)-desaturase"		
	/protein_id="AAD04034.1"		
	/db_xref="GI:4140398"		
	/translation="MEDYLFQVETSFYNRLVLGTFMPSSWGWGFLPHMLOGWLNRYI GGVLLFISGLFMCFIYIHUKRNYIIPDAIPSNKMLLQISVAMKAPWTCAFPSLS EYMIENGWTKFARISDVGLSWLYIAAYLVIVFEGIYMMHMLHDIPKLYKYLHAT HHYINQNTKSPAGVAFHPLDGLIQAVPHVALLVPMHFHSITIALIFLEALVTANI HDCIHGKVFVPMVGAGYHTIHRRYRHNHYGTIIMDMWFGTLRDPVEEDAKKM"		
CDS			
	310..239	239 c 245 g 361 t	
BASE COUNT			
ORIGIN			
	Query Match	13.2%	Score 186.2; DB 8; Length 1155;
	Best Local Similarity	72.2%	Pred. No. 3.2e-30;
	Matches	242; Conservative	0; Mismatches 93; Indels 0; Gaps 0;
QY	578	GTCTGTTTCGCTTCAGATGCAATTCCTACATAAGCGTATGCGTTTCCAAATGTTCTG	637
Db	304	GTATATTTCTCTAAAGATGCCATCCCTTCTAACAAAGCAATGCTCTTGCAATATCATGA	363
QY	638	GCAATGIAGCGCTATGCCATGGTACACATCTCTTCCCAACTGCTCTCCGAGAGTATGATTGA	697
Db	364	GCCATGIAGCGGATGCCATGGTACTGTCGCCCTCCATCATTCTCTGAGTACATGATTGA	423
QY	698	CGTGGTGGACAAAAGTTTTTGGCTAGCATACAGCAATTCGCGCTCGATTCTGTATTTGTT	757
Db	424	AATGGGAAGCAAAAATTTTTGCAAGAATAAGTATGTTGGATGGCTTTTCCTGATTATC	483

QY 698 CGTGGTGGACCAAAATGTTTGTCTAGCATAGACGAATTCGGCTGGGATCTGATATTTGTT 757
DB 448 AATGGGTGGACAAAATGTTTCTCAAGAAATAGCGAGTTGGATGGCTTCCCTACCTATT 507
QY 758 TACATGCCATCTATCTCTTTTCTGTTGAGTTTGGTATTATTATTTGGATGACAGAGCTT 817
DB 508 TATATGCCAGTTTATTGTTGATAGTAGAATTCGGCAATTACTGGATGATCGCGAGTTG 567
QY 818 CATGACATTAAGCTCTCTATAGTATCTCCATGCCACCATCATATCTACAAACAGCAG 877
DB 568 CATGACATTAAGCTCTCTACAAATATCTCCATGCTACACATCATATTTACAAACAGCA 627
QY 878 AATACACTCTCTCCATTCGCGGTAAGTGTTCFA 912
DB 628 AATACACTCTCCCACTTGTGTTGGGTTCCA 662

RESULT 9
AX155063
LOCUS 1381 bp DNA linear PAT 22-JUN-2001
DEFINITION Sequence 109 from Patent WO0138484.
ACCESSION AX155063
VERSION AX155063.1 GI:14536617
KEYWORDS Physcomitrella patens.
SOURCE Physcomitrella patens
ORGANISM Physcomitrella patens
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Bryophyta;
Bryopsida; Funariidae; Funariales; Funariaceae; Physcomitrella.
REFERENCE 1 (bases 1 to 1381)
AUTHORS Lerchl, J., Renz, A., Ehrhardt, T., Reindl, A., Cirpus, P., Bischoff, F.,
Frank, M., Freund, A., Duwenig, E., Schmidt, R.M. and Reski, R.
TITLE Moss genes from Physcomitrella patens encoding proteins involved in
the synthesis of polyunsaturated fatty acids and lipids
JOURNAL Patent: WO 0138484-A 109 31-MAY-2001;
BASF Plant Science GmbH (DE)
FEATURES
source location/Qualifiers
1. .1381
/organism="Physcomitrella patens"
/db_xref="taxon:3218"
40. .1041
/note="PP004064012R"
/codon_start=1
/protein_id="CAC42667.1"
/db_xref="GI:14536618"
/translation="MASRGAVNMVCAIAIVLMVMSLSICMSADVEVVNASFSSVVG
GAKTSGVVPANGSEYLAIFVEETRWYNDLVGLPWPSSVRDSSIPHTLOTWLRNV
AGMLYFVSGGLWCLYVYSKGEHEFPAGDIPAKPIMLQIWTWKAMPVYTGCLPLS
EYMERGWTCKAREDEDVGLYVGLVIAVLAVVERGIYWMHRELIDIKPLKHLAT
HHIYNKQNTLSFPAGLAFHPIDGILQACPHVIALFLPMHFFTHVELLFCGVTWTNI
HDCIDGNVNGIMGAGFHTTHHTYRHYNGHYTFVMDWLFGLTRDPYERRKATARVKS"

BASE COUNT 323 a 331 c 351 g 376 t

Query Match 11.6%; Score 164.2; DB 6; Length 1381;
Best Local Similarity 68.7%; Pred. No. 1.8e-25;
Matches 226; Conservative 0; Mismatches 103; Indels 0; Gaps 0;

QY 584 TTTCCTTCAGATCAATTCCTACAAATAAGGCTATCGTTCGAAATGTTGGCAATG 643
DB 421 TTCCCTGCGAGTGACATACCGCGAAGGAGCCCAATATGCTCCAAATCTGGGTAATG 480
QY 644 AAGGCTATGCCATGCTACTCTCTTCCCACTGTCTCCGAGAGTATGATTGAACGTGGT 703
DB 481 AAGGCTATGCCATGCTACTATACAGAGCTTCCCACTCTCTCCGAAATATATGATTGAGCGGGG 540
QY 704 TGGACCAAAATGTTTCTAGCATAGACGAATTCGGCTGGATCTGATTTGTTTACATC 763
DB 541 TGGACCAAAATGTTTCTAGCATAGACGAATTCGGCTGGATCTGATTTGTTTACATC 600
QY 764 GCATCTATCTGTTTCTGAGTTGGTATTATTATTTGGATGACAGAGAGCTTCATGAC 823
DB 601 ATCGCCTACTTGGCACTGGTGAGTTGGTATTATCTATTGGATGACAGAGAGCTTCAGAT 560

QY 824 ATTAAGCCTCTCTATAGTATCTCCATGCCACCATCATATCTACAAACAGCAGATACA 883
DB 661 ATTAAGCCTTTGTACAAACATCTGCTATGCCACCATCATCTACAAACAGCAGACG 720
QY 884 CTTCTCTCCATTTGCCCGGTAAAGTCTTTTCA 912
DB 721 CTATACCCGTTTGCAGGTTTGGCGTTCCA 749

RESULT 10
G71681/c
LOCUS 293 bp DNA linear STS 08-JUN-2001
DEFINITION A85715834FMO17 maize leaf DNA may be STS genomic, sequence tagged
site.
ACCESSION G71681
VERSION G71681.1 GI:14333366
KEYWORDS STS.
SOURCE Zea mays.
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 293)
AUTHORS Yang, Y.J., Guo, L., Ashlock, D.A., Wen, T.J. and Schnable, P.S.
TITLE 3' UTR sequences of maize genes
JOURNAL Unpublished (2001)
COMMENT Contact: Schnable, P.S.
Schnable laboratory
Iowa State University
G405 Agronomy Hall, Ames, IA 50011, USA
Tel: 515-294-0975
Fax: 515-294-2299
Email: schnable@iastate.edu
Primer A: CTAAATAGCAGACCAAGCC
Primer B: GATGTATCTGCAAGCG
PCR Profile:
Denaturation: 94 degrees C for 30 seconds
Annealing: 60 degrees C for 45 seconds
Polymerization: 72 degrees C for 90 seconds
PCR cycles: 31
Thermal cycler: Perkin Elmer TC
Protocol:
Template: 10-20 ng
Primer: each 0.5 uM
dNTPs: each 200 uM
Taq Polymerase: 0.05 units/uL
Total vol: 20 uL
Buffer:
MgCl2: 2 mM
KCl: 50 mM
Tris-HCl: 20 mM
pH: 8.4.
Location/Qualifiers
1. .293
/organism="Zea mays"
/strain="DE811"
/db_xref="taxon:4577"
/clone_lib="maize leaf DNA"
/note="PCR products amplified from genomic DNA"

BASE COUNT 77 a 51 c 81 g 84 t

Query Match 6.6%; Score 92.6; DB 11; Length 293;
Best Local Similarity 71.3%; Pred. No. 8.6e-10;
Matches 122; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

QY 1240 GTTTTGTGTTACAGGGCTTGCATTTCCACCCAGTAGCGGATACTTTCAGGCTGTACCGCA 1299
DB 226 GTCTCCATTTGCTGGACTTGCATTTCCACCCAGTAGGATGTTCTTCTGCAAGCAATACCA 167

QY 1300 TGTGATAACGCTGTTTATAGTCCCAATTCATTTCACAACTCATATAGGCTCTTTGTTTCAT 1359
 Db 166 TGTGTTTCACGCTCTCTCTCTTTTCCAAACGCACTTCAGGACACATATGCTCTCTGTTCTT 107
 QY 1360 GGAAGCGTATGGAGCGGCGAATCCATCCATGATGCTATCCATGCGCAACATCTG 1410
 Db 106 AGAGGGTGTGGGACAAACATCCAGGACTGCTATTCATGCGCAAGATATG 56

RESULT 11

G71612/c
 LOCUS G71612 310 bp DNA linear STS 08-JUN-2001
 DEFINITION A6298234:Mo17 maize leaf DNA Zea mays STS genomic, sequence tagged site.
 ACCESSION G71612
 VERSION G71612.1 GI:14333297
 KEYWORDS STS.
 SOURCE Zea mays.
 ORGANISM Zea mays.
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Zea.
 REFERENCE 1 (bases 1 to 310)
 AUTHORS Yang, Y.J., Guo, L., Ashlock, D.A., Wen, T.J. and Schnable, P.S.
 TITLE 3' UTR sequences of maize genes
 JOURNAL Unpublished (2001)
 COMMENT Contact: Schnable, P.S.
 Schnable laboratory
 Iowa State University
 G405 Agronomy Hall, Ames, IA 50011, USA
 Tel: 515-294-0975
 Fax: 515-294-2299
 Email: schnable@iastate.edu
 Primer A: CTAATAGACACCAAGCC
 Primer B: GATGGATCTCGAAGCG
 PCR Profile:
 Denaturation: 94 degrees C for 30 seconds
 Annealing: 60 degrees C for 45 seconds
 Polymerization: 72 degrees C for 90 seconds
 PCR cycles: 31
 Thermal cycler: Perkin Elmer TC

Protocol:

Template: 10-20 ng
 Primer: each 0.5 uM
 dNTPs: each 200 uM
 Taq Polymerase: 0.05 units/ul
 Total vol: 20 ul

Buffer:

MgCl2: 2 mM
 KCl: 50 mM
 Tris-HCl: 20 mM
 pH: 8.4.

Location/Qualifiers

1..310
 /organism="Zea mays"
 /strain="DE811"
 /db_xref="taxon:4577"
 /clone_lib="maize leaf DNA"
 /note="PCR products amplified from genomic DNA"
 <1..>310

STS

BASE COUNT 80 c 87 g 62 t

ORIGIN

Query Match 4.7%; Score 65.8; DB 11; Length 310;
 Best Local Similarity 70.4%; Pred. No. 0.00056;
 Matches 88; Conservative 0; Mismatches 37; Indels 0; Gaps 0;
 QY 1286 CAGGCTTACGCGCATGTAGAGCGCTGTTTATAGTCCCAATTCATTTCACAACTCATATA 1345
 Db 309 CAAGCAATGGCGCATGTGTGCGCTCTTCTCTCCCAAGCGCACTTCAGGACGCATC 250

QY 1346 GGTCTTTTGTTCATGGAAGCGATATGAGCGGAACATCCATGACTGCATCCATCCATGCAAC 1405
 Db 249 GCTCTCGTGTCTTGGAGGGGTGGAGGACAAACATCCAGCACTGCATCCACGCGAAG 190
 QY 1406 ATCTG 1410
 Db 189 TTATG 185

RESULT 12

CNS0180K
 LOCUS CNS0180K 840 bp mRNA linear PLN 02-SEP-1999
 DEFINITION Botrytis cinerea strain T4 CDNA library under conditions of nitrogen deprivation.
 ACCESSION AL110675
 VERSION AL110675.1 GI:5824962
 KEYWORDS CDNA library; nitrogen deprivation.
 SOURCE Botryotinia fuckeliana.
 ORGANISM Botryotinia fuckeliana.
 Eukaryota; Fungi; Ascomycota; Pezizomycotina; Leotiomycetes; Helotiales; Sclerotiniaceae; Botryotinia.
 REFERENCE 1 (bases 1 to 840)
 AUTHORS Bitton, F., Lewis, C., Fortini, D., Pradier, J.M. and Brygoo, Y.
 TITLE Direct Submission
 JOURNAL Submitted (01-SEP-1999) Phytopathologie, INRA, route de St Cyr, 78026 Versailles, France
 REFERENCE 2 (bases 1 to 840)
 Genoscope.
 Direct Submission
 Submitted (01-SEP-1999) Genoscope - Centre National de Sequencage; CP 5706 91057 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)

COMMENT The cDNA library to be analyzed within the framework of this project was created using a Botrytis cinerea strain which was grown under conditions of nitrogen deprivation, which is the normal situation for B. cinerea during its development on its host plant. The library was produced in an oriented direction, in the pBSII vector.

FEATURES

Source
 1..840
 /organism="Botryotinia fuckeliana"
 /strain="T4"
 /db_xref="taxon:40559"
 /note="Genoscope sequence ID : W33E091"

BASE COUNT 151 a 23 c 79 g 539 t 48 others

ORIGIN

Query Match 4.1%; Score 57.6; DB 8; Length 840;
 Best Local Similarity 46.3%; Pred. No. 0.028;
 Matches 186; Conservative 0; Mismatches 216; Indels 0; Gaps 0;
 QY 235 TTTCACCTTTTCATTCACATATGCTTAATCGCTTCTATGTTATCGATTTTCAATTTAA 294
 Db 249 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 308
 QY 295 GGAAGAGGGTTTCTTCGTGCTACTGTACAGTAAATTTGGATTTGATGCGATGTTATGTT 354
 Db 309 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 368
 QY 355 TGCATTTATGATTAATTTGTGCATATCTCCATCTCAAGGATGAACAGTTAGTGGCTTA 414
 Db 369 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 428
 QY 415 TATAAGTTTGTGCAACCAATGAGAAGCTGTACATCTTGAAGTTGAATTTCTACTGTG 474
 Db 429 TTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 488
 QY 475 CCATTTAAGTCCACTTAAATTTGTTTGAAGTGTATGCTACTTCTAGACACATCTTT 534
 Db 489 TATTATTTATTAATTAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 548
 QY 535 TTCTGCTTCTCTGAGACTCTGCTCTAGTTTGAATCTTTTGGTCTGTTTGGTCTCAGA 594

[illegible]

Db 10196 TTCGCCGATCGATTGAATATACNCCGCCCTGTNNNNNNNNNNNNNNNNNN 10137
QY 195 TTACCTTAAATCAACGGTTTACCTCCCAAGGCTCGACCTTTCACCTTTTGTATTCACCTA 254
Db 10136 NNNNNNTNN 10077
QY 255 TTGCTTAATCGCTTTCTAAGTATGATGATTTTCAATTTAAGGAAGAGGGTTTCTTCGGTGT 314
Db 10076 TTTTTCCTTTTCTTTTGTGTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTNNN 10017
QY 315 ACTGACACGTAATTTGGATTTGATGCTGATGATGATGATGATGATGATGATGATGATGAT 374
Db 10016 NTTTTNNNTTTTTNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 9957
QY 375 GCATATTTCCATCTAAGGATTAACACAGTTAGTGGCTATATATAAGTTTGTGCAACCA 434
Db 9956 TNNTTTTNTTTTTNTTTTTTTTTNNNNNNNNNNNNNNNNNNNNNNNNNNNN 9897
QY 435 ATGAGAACTCGTACATCTTTGAAGTGTGAATTTTCTACCTGCAATTTAAGTCCACTAAAT 494
Db 9896 TTTTNTN'TTTTTTTTTNTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 9837
QY 495 TGTGTGTGAAGTGTGCTACTTTCAGACACATCTTTCTCTCTCTCTGAGACCT 554
Db 9836 NTTTTTTTNTTT 9777
QY 555 GTCTTAGTTGAAATCTTTTTTGGTCTGTTTGGCTTCAGAT 595
Db 9776 TT 9736

RESULT 15
AC010772/c
LOCUS
DEFINITION Homo sapiens chromosome 15 clone 318_H_11 map 15, LOW-PASS SEQUENCE
SAMPLING.
ACCESSION AC010772
VERSION AC010772.1 GI:5917986
KEYWORDS HTG; HTGS-PHASE0.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Birken, B., Linton, L., Nusbaum, C. and Lander, E.
Homo sapiens chromosome 15, clone 318_H_11
Unpublished
2 (bases 1 to 52359)
Birken, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, M.,
Baldwin, C., Barna, N., Beckerly, R., Boguslavsky, L., Boukhalter, B.,
Brown, A., Castle, A., Collings, M., Collins, S., Collymore, A.,
Cooke, P., Dearellano, K., Dewar, K., Domino, M., Donelan, L., Doyle, M.,
Ferrelira, P., FitzHugh, W., Forrest, C., Funke, R., Gage, D.,
Galagan, J., Gardyna, S., Grant, G., Hagos, B., Heaford, A., Horton, L.,
Howard, J., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J.,
Lehoczky, J., Lieu, C., Locke, K., Macdonald, P., Marquis, N.,
McEwan, P., McGurk, A., McKernan, K., McLaughlin, J., Meldrum, J.,
Morrow, J., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P.,
Peterson, K., Pollara, V., Riley, R., Roy, A., Santos, R., Severy, P.,
Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,
Testaye, S., Tirrell, A., Vassiliev, H., Vo, A., Wheeler, J., Wu, X.,
Wyman, D., Ye, W. J., Zimmer, A. and Zody, M.
Direct Submission
Submitted (22-SEP-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
All repeats were identified using RepeatMasker: Smit, A. F. A. &
Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>.
* NOTE: this record contains 57 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for

* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.
* 1
* 914: contig of 914 bp in length
* gap of unknown length
* 915 1826: contig of 912 bp in length
* gap of unknown length
* 1827 2677: contig of 851 bp in length
* gap of unknown length
* 2678 3628: contig of 951 bp in length
* gap of unknown length
* 3629 4580: contig of 952 bp in length
* gap of unknown length
* 4581 5603: contig of 1023 bp in length
* gap of unknown length
* 5604 6470: contig of 867 bp in length
* gap of unknown length
* 6471 7393: contig of 923 bp in length
* gap of unknown length
* 7394 8268: contig of 875 bp in length
* gap of unknown length
* 8269 9149: contig of 881 bp in length
* gap of unknown length
* 9150 10088: contig of 939 bp in length
* gap of unknown length
* 10089 10983: contig of 895 bp in length
* gap of unknown length
* 10984 11893: contig of 910 bp in length
* gap of unknown length
* 11894 12759: contig of 866 bp in length
* gap of unknown length
* 12760 13692: contig of 933 bp in length
* gap of unknown length
* 13693 14592: contig of 900 bp in length
* gap of unknown length
* 14593 15485: contig of 893 bp in length
* gap of unknown length
* 15486 16433: contig of 948 bp in length
* gap of unknown length
* 16434 17282: contig of 849 bp in length
* gap of unknown length
* 17283 18170: contig of 888 bp in length
* gap of unknown length
* 18171 19042: contig of 872 bp in length
* gap of unknown length
* 19043 19980: contig of 938 bp in length
* gap of unknown length
* 19981 20951: contig of 971 bp in length
* gap of unknown length
* 20952 21821: contig of 870 bp in length
* gap of unknown length
* 21822 22778: contig of 957 bp in length
* gap of unknown length
* 22779 23762: contig of 984 bp in length
* gap of unknown length
* 23763 24711: contig of 949 bp in length
* gap of unknown length
* 24712 25619: contig of 908 bp in length
* gap of unknown length
* 25620 26561: contig of 942 bp in length
* gap of unknown length
* 26562 27491: contig of 930 bp in length
* gap of unknown length
* 27492 28397: contig of 906 bp in length
* gap of unknown length
* 28398 29331: contig of 934 bp in length
* gap of unknown length
* 29332 30263: contig of 932 bp in length
* gap of unknown length
* 30264 31248: contig of 985 bp in length

* 31249 gap of unknown length
* 32162: contig of 914 bp in length
* gap of unknown length
* 33089: contig of 927 bp in length
* gap of unknown length
* 34038: contig of 949 bp in length
* gap of unknown length
* 34936: contig of 898 bp in length
* gap of unknown length
* 34937: contig of 1036 bp in length
* gap of unknown length
* 35973: contig of 894 bp in length
* gap of unknown length
* 36867: contig of 911 bp in length
* gap of unknown length
* 37778: contig of 905 bp in length
* gap of unknown length
* 38683: contig of 900 bp in length
* gap of unknown length
* 39583: contig of 871 bp in length
* gap of unknown length
* 40454: contig of 922 bp in length
* gap of unknown length
* 41376: contig of 920 bp in length
* gap of unknown length
* 42296: contig of 933 bp in length
* gap of unknown length
* 43229: contig of 922 bp in length
* gap of unknown length
* 44151: contig of 920 bp in length
* gap of unknown length
* 45071: contig of 934 bp in length
* gap of unknown length
* 46005: contig of 870 bp in length
* gap of unknown length
* 46875: contig of 845 bp in length
* gap of unknown length
* 47720: contig of 933 bp in length
* gap of unknown length
* 48653: contig of 912 bp in length
* gap of unknown length
* 49565: contig of 912 bp in length
* gap of unknown length
* 50477: contig of 905 bp in length
* gap of unknown length
* 51382: contig of 978 bp in length.
* Location/Qualifiers
1. .52359
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="15"
/map="15"
/clone="318_H_11"
/clone_lib="RPC1-11 Human Male BAC"
BASE COUNT 14555 a 7384 c 13014 g 7899 t 9507 others
ORIGIN

FEATURES
source

Query Match 4.0%; Score 55.8; DB 2; Length 52359;
Best Local Similarity 34.5%; Pred. No. 0.034;
Matches 138; Conservative 0; Mismatches 262; Indels 0; Gaps 0;

QY 235 TTTCACCTTTGTATTCACATTCGCTTAATCGCTTTCTATGTTATCGATTTTTCATTTAA 294
Db 48642 TTTTNTNNNT 48583
QY 295 GGAGAGGGTTTCCTGCTACTGTACAGTAATTTGGATTTGATGGGATGTCATGTT 354
Db 48582 TTTTNNNTNN 48523
QY 355 TGCATTTATGATTTATTTGTCATATTCCTCAATCTAAGGATTAAGACAGTACTGGCTTA 414
Db 48522 TTTTNNNN 48463

QY 415 TATAAGTTTTTGTGCAACCAATGAGAAGTCGTACATCTTTTGAAGTTGAATTTTCTACTTGG 474
Db 48462 TTTTNTTTT 48403
QY 475 CCATTTAAGTCCACCTAAATTTGTTCTTGAAGTGATTGCTACTTTTCACACACATTCITT 534
Db 48402 TTTTNTTTT 48343
QY 535 TTCTGCTTCTCTGAGACTCTGCTTAAAGTTTGAATCTTTTGGTCTGTTTGGCTTCAGA 594
Db 48342 TTTTNTTTT 48283
QY 595 TGAATTCCTACAATAAAGGCTATGCGTTTGCAAAATGTTT 634
Db 48282 NNT 48243

Search completed: June 6, 2003, 15:14:08
Job time : 4084.7 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 6, 2003, 10:02:23 ; Search time 520.302 Seconds
(without alignments)
10068.203 Million cell updates/sec

Title: US-09-775-879-20_COPY_143_322
Perfect score: 180
Sequence: 1 atgucggcgataatgctta.....tcattcgcgtttctctctgg 180

Scoring table: IDENT(TY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054610 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb.ba.*
- 2: gb.htg.*
- 3: gb.in.*
- 4: gb.om.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.ph.*
- 8: gb.pl.*
- 9: gb.pr.*
- 10: gb.ro.*
- 11: gb.sts.*
- 12: gb.sy.*
- 13: gb.un.*
- 14: gb.vi.*
- 15: em.ba.*
- 16: em.fun.*
- 17: em.hum.*
- 18: em.in.*
- 19: em.mu.*
- 20: em.om.*
- 21: em.or.*
- 22: em.ov.*
- 23: em.pat.*
- 24: em.ph.*
- 25: em.pl.*
- 26: em.ro.*
- 27: em.sts.*
- 28: em.un.*
- 29: em.vi.*
- 30: em.htg.hum.*
- 31: em.htg.inv.*
- 32: em.htg.other.*
- 33: em.htg.mus.*
- 34: em.htg.pln.*
- 35: em.htg.rod.*
- 36: em.htg.mam.*
- 37: em.htg.vrt.*
- 38: em.sy.*
- 39: em.htgo.hum.*
- 40: em.htgo.mus.*
- 41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	180	100.0	1163	8	AY088527	Arabidops
2	180	100.0	2294	8	AF105034	Arabidops
3	180	100.0	3004	8	AF069468	Arabidops
4	180	100.0	103904	8	ATAC021640	Arabidops
5	178.4	99.1	1175	8	AF081794	A.thaliana
6	82	45.6	1155	8	AF081794	A.thaliana
7	67.6	37.6	179428	8	AF003214	Oryza sat
8	61.4	34.1	1148	8	AF099969	Nicotiana
9	54.8	30.4	1381	6	AX155063	Sequence
10	35.8	19.9	162212	2	AC113675	Rattus no
11	35.8	19.9	174556	2	AC107404	Rattus no
12	34.8	19.3	116788	9	AC007247	Homo sapi
13	33.2	18.4	186537	2	AC116209	Rattus no
14	33	18.3	190998	9	AC019181	Homo sapi
15	32.6	18.1	10393	1	U32814	Haemophilus
16	32.6	18.1	108795	2	AC124894	Rattus no
17	32.6	18.1	156491	9	AL591720	Human DNA
18	32.6	18.1	159914	9	AC027553	Homo sapi
19	32.6	18.1	168974	2	AC021299	Homo sapi
20	32.6	18.1	185822	2	AC073554	Homo sapi
21	32.6	18.1	196832	2	AL772211	Mus muscu
22	32.6	18.1	225957	9	AC069175	Homo sapi
23	32.4	18.0	31975	8	SPCC553	AL023704 S.pombe c
24	32.4	18.0	42776	3	CEF1504	280344 Caenorhabdi.
25	32.4	18.0	100590	8	OSJN00191	AL662991 Oryza sat
26	32.2	17.9	254752	2	AC127336	Mus muscu
27	32	17.8	120416	9	AC093915	Homo sapi
28	32	17.8	155223	9	AP001528	Homo sapi
29	32	17.8	175376	2	AC108223	Oryza sat
30	32	17.8	179491	2	AC020673	Homo sapi
31	32	17.8	184439	2	AP003085	Homo sapi
32	31.8	17.7	80568	9	AC002420	Human Chr
33	31.8	17.7	107708	9	AL139803	Human DNA
34	31.8	17.7	158073	9	AC069298	Homo sapi
35	31.8	17.7	164781	9	AC092440	Homo sapi
36	31.8	17.7	171031	2	AC104020	Homo sapi
37	31.8	17.7	174649	9	AC103586	Homo sapi
38	31.8	17.7	176030	2	AC100735	Mus muscu
39	31.8	17.7	180126	2	AC022074	Homo sapi
40	31.8	17.7	210000	2	AC006295	Homo sapi
41	31.8	17.7	232065	2	AC069425	Homo sapi
42	31.6	17.6	62088	2	AC104418	Homo sapi
43	31.6	17.6	66017	2	AC124262	Homo sapi
44	31.6	17.6	82642	9	AC103745	Homo sapi
45	31.6	17.6	156741	2	AC131312	Homo sapi

ALIGNMENTS

RESULT 1
AY088527
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS

AY088527
Arabidopsis thaliana clone 7564 mRNA, complete sequence.
AY088527
AY088527.1 GI:21407301
FLI_CDNA.
thale cress.
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 1163)
Haas,B.J., Volfovsky,N., Town,C.D., Troukhan,M., Alexandrov,N.,
Feldmann,K.A., Flavell,R.B., White,O. and Salzberg,S.L.

1163 bp mRNA linear PLN 26-JUN-2002

TITLE	Full-length messenger RNA sequences greatly improve genome annotation
JOURNAL	Genome Biol. (2002) In press
REFERENCE	2 (bases 1 to 1163)
AUTHORS	Prover, V., Troukhan, M., Alexandrov, N., Lu, Y.-P., Flavell, R. and Feldmann, K.
TITLE	Full-length cDNA from Arabidopsis thaliana
JOURNAL	Unpublished
REFERENCE	3 (bases 1 to 1163)
AUTHORS	Prover, V., Troukhan, M., Alexandrov, N., Lu, Y.-P., Flavell, R. and Feldmann, K.
TITLE	Direct Submission
JOURNAL	Submitted (11-MAR-2002) Ceres, Inc. 3007 Malibu Canyon Road, Malibu, CA 90265, USA
COMMENT	This clone sequence is one of 5,000 Ceres full-length cDNAs made available to TIGR and Genbank. The following quality assessment of this set was done by comparison with known proteins: two percent of the clones are estimated to be 5'-truncated; less than one percent are 3'-truncated; approximately two percent represent alternative splice variants, including unspliced introns and spliced exons; one percent may contain premature stop codons; five percent may have frame shifts in a coding region. A sequence is considered to be 5'-truncated if it lacks the translation initiation start (ATG). A sequence is considered to be 3'-truncated if it lacks the C-terminal end of the encoded protein. Please note that these cDNA sequences are derived from the Ws or Laer ecotypes and therefore may contain polymorphisms when compared to sequences from Col-0. Genet carried out the library production and sequencing of the full-length clones, Ceres, Inc. carried out the clustering of the 5' sequences, selection of clones, and sequence assembly.
FEATURES	Location/Qualifiers
source	1..1163 /organism="Arabidopsis thaliana" /db_xref="taxon:3702" /clone="7564"
CDS	146..991 /codon_start=1 /product="sterol-C5-desaturase" /protein_id="AAM66060.1" /db_xref="GI:21594971" /translation="MAADNAYLMQFVDETSFYNRIVLSHLLPANLWELPHELOTWLR NYLAGTLLFISGLWCIFYIYKINYLKPKDAIPKAMRLQMFVAMKAMPYLLP TVSEMIKRGNTKSPASIDFGWILFYIAYLVFVEGIYMHRELHDKPLIYKIL HATHHLYKQNTLSFAGLAFHPVDGILQAVPHVIALFIVPHFTTHIGLLEFMEAIWT ANIHDICIGNIPVMGAGYVTHHTTHYKHYGHTIWMDFGSLRDLLEEDDNKDS FKKAE"
BASE COUNT	291 a 263 c 223 g 384 t 2 others
ORIGIN	
Query Match	100.0%; Score 180; DB 8; Length 1163;
Best Local Similarity	100.0%; Pred. No. 1.6e-44;
Matches 180; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
QY	1 ATGGCGCGGATAATCCTTATCTATCTGATGACAGTTGTTGACGAACCTCTTTTACACCGA 60
Db	146 ATGGCGCGGATAATCCTTATCTATCTGATGACAGTTGTTGACGAACCTCTTTTACACCGA 205
QY	61 ATCGTTCTGAGTCATCTTTTGGCGGGAATCTATGGGAACCTTACTCTATCTCTCTCCAG 120
Db	206 ATCGTTCTGAGTCATCTTTTGGCGGGAATCTATGGGAACCTTACTCTATCTCTCTCCAG 265
QY	121 ACATGCTCGGAATACCTTCGCGGGAACCTTACTACTTACTTATCTATCTCTCTCTCTCG 180
Db	266 ACATGCTCGGAATACCTTCGCGGGAACCTTACTACTTACTTATCTATCTCTCTCTCTCG 325
RESULT 2	
AF105034	
LOCUS	AF105034 2294 bp DNA PLN 10-JUN-1999
DEFINITION	Arabidopsis thaliana delta7 sterol C-5 desaturase (STEL) gene, complete cds.
ACCESSION	AF105034
VERSION	AF105034.1 GI:5031218

KEYWORDS	Arabidopsis thaliana.
SOURCE	Arabidopsis thaliana
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
REFERENCE	1 (bases 1 to 2294)
AUTHORS	Choe, S., Noguchi, T., Fujioka, S., Takatsuto, S., Tissier, C.P., Gregory, B.D., Ross, A.S., Tanaka, A., Yoshida, S., Tax, F.E. and Feldmann, K.A.
TITLE	The Arabidopsis dwf7/stel mutant is defective in the delta7 sterol C-5 desaturation step leading to brassinosteroid biosynthesis
JOURNAL	Plant Cell 11 (2), 207-221 (1999)
MEDLINE	99128169
PUBMED	9927639
REFERENCE	2 (bases 1 to 2294)
AUTHORS	Choe, S., Tanaka, A., Gregory, B.D. and Feldmann, K.A.
TITLE	Direct Submission
JOURNAL	Submitted (06-NOV-1998) Plant Sciences, The University of Arizona, Forbes Hall 303, Tucson, AZ 85721, USA
FEATURES	Location/Qualifiers
source	1..2294 /organism="Arabidopsis thaliana" /cultivar="Ws-2 (Wassilewskija-2)" /db_xref="taxon:3702" /chromosome="III" /map="2.2" 43..1889 /gene="STEL" /note="DWARF7"
gene	43..1889 /gene="STEL" /note="sterol/brassinosteroid biosynthetic enzyme"
mRNA	Join(43..368,736..1041,1396..1889) /gene="STEL"
CDS	/product="delta7 sterol C-5 desaturase" Join(143..368,736..1041,1396..1709) /gene="STEL" /note="sterol/brassinosteroid biosynthetic enzyme"
variation	322 /gene="STEL" /note="dwf7-2 (stel-3) allele" /replace="a"
variation	1552 /gene="STEL" /note="dwf7-1 (stel-2) allele" /replace="a"
BASE COUNT	585 a 469 c 417 g 823 t
ORIGIN	
Query Match	100.0%; Score 180; DB 8; Length 2294;
Best Local Similarity	100.0%; Pred. No. 1.7e-44;
Matches 180; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
QY	1 ATGGCGCGGATAATCCTTATCTATCTGATGACAGTTGTTGACGAACCTCTTTTACACCGA 60
Db	143 ATGGCGCGGATAATCCTTATCTATCTGATGACAGTTGTTGACGAACCTCTTTTACACCGA 202
QY	61 ATCGTTCTGAGTCATCTTTTGGCGGGAATCTATGGGAACCTTACTCTATCTCTCTCCAG 120
Db	203 ATCGTTCTGAGTCATCTTTTGGCGGGAATCTATGGGAACCTTACTCTATCTCTCTCCAG 262
QY	121 ACATGCTCGGAATACCTTCGCGGGAACCTTACTACTTACTTATCTATCTCTCTCTCTCG 180
Db	263 ACATGCTCGGAATACCTTCGCGGGAACCTTACTACTTACTTATCTATCTCTCTCTCTCG 322

RESULT 3	AF069468	3004 bp	DNA	linear	PLN 19-MAY-1999
LOCUS	Arabidopsis thaliana sterol-C5-desaturase gene, complete cds.				
DEFINITION	AF069468				
ACCESSION	AF069468				
VERSION	AF069468.1	GI:4234767			
KEYWORDS					
SOURCE	Arabidopsis thaliana.				
ORGANISM	Arabidopsis thaliana				
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;				
	Rosidae; euroids II; Brassicales; Brassicaceae; Arabidopsis.				
REFERENCE	1 (bases 1 to 3004)				
AUTHORS	Husselstein, T., Schaller, H., Gachotte, D. and Benveniste, P.				
TITLE	Delta7-sterol-C5-desaturase: molecular characterization and functional expression of wild-type and mutant alleles				
JOURNAL	Plant Mol. Biol. 39 (5), 891-906 (1999)				
MEDLINE	99273992				
PUBMED	10344195				
REFERENCE	2 (bases 1 to 3004)				
AUTHORS	Husselstein, T., Schaller, H., Gachotte, D. and Benveniste, P.				
TITLE	Direct Submission				
JOURNAL	Submitted (30-MAY-1998) IBMP, 28, rue Goethe, Strasbourg 67083, France				
FEATURES	source				
	location/Qualifiers				
	1..3004				
	/organism="Arabidopsis thaliana"				
	/cultivar="columbia"				
	/db_xref="taxon:3702"				
	/tissue_type="leaves"				
	join(1242..1467,1836..2141,2496..2809)				
	/codon_start=1				
	/product="sterol-C5-desaturase"				
	/protein_id="AAD12944.1"				
	/db_xref="GI:4234768"				
	/translation="MAADNAYLMQFVDTSEYNRIVLSHLLPANLWELPLFLQTWLR				
	NYLAGTLFTSGFLWCFIYLYLKVLPKDAIPTKAMLOMFVAMKMPWTLPL				
	TVSMTGRTGKTFASIGFEGWILYFYIAYLVFVEFYWMHRLHDKPLYKYL				
	HATHIYNKNTLSPFAGLAPHPVDGILQAPVHVALFIVPIHTHTGLLEFMEAIWT				
	ANIHDHIGNLWPVWGAGYHTIHTHTYKHNYGHTIYNDWMFGSLRDLLEEDNDKDS				
	FKAAE"				
CDS	865 .. 577 c. 490 g 1072 t				
BASE COUNT					
ORIGIN					
	Query Match	100.08;	Score 180;	DB 8;	Length 3004;
	Best Local Similarity	100.08;	Pred. No. 1.7e-44;		
	Matches 180;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	ATGGCG:CGGATAATGCTTATCTGATGTCAGTTTGTGACGAACCTCTTTTACACCGA	60		
Db	1242	ATGGCG:CGGATAATGCTTATCTGATGTCAGTTTGTGACGAACCTCTTTTACACCGA	1301		
Qy	61	ATCGTT:TGAGTCATCTTTTCGCGGCGGAATCTATGGGAACCTTACCTCATTTTCTCCAG	120		
Db	1302	ATCGTT:TGAGTCATCTTTTCGCGGCGGAATCTATGGGAACCTTACCTCATTTTCTCCAG	1361		
Qy	121	ACATGG:TCCGAATACCTCGCGGCGGAACCTACTACTACTATCTCTCGGTTTCTCTGG	180		
Db	1362	ACATGG:TCCGAATACCTCGCGGCGGAACCTACTACTACTATCTCTCTCTGG	1421		
RESULT 4	ATAC021640	103904 bp	DNA	linear	PLN 24-JAN-2001
LOCUS	Arabidopsis thaliana chromosome III BAC F16B3 genomic sequence, complete sequence.				
DEFINITION	ATAC021640				
ACCESSION	ATAC021640				
VERSION	ATAC021640.7	GI:12408747			
KEYWORDS	HTG.				
SOURCE	Arabidopsis thaliana.				
ORGANISM	Arabidopsis thaliana				
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;				
	Rosidae; euroids II; Brassicales; Brassicaceae; Arabidopsis.				
REFERENCE	1 (bases 1 to 103904)				
AUTHORS	Lin, X., Kaul, S., Town, C.D., Benito, M.-I., Creasy, T.H., Haas, B., Wu, D., Rensing, C.M., Koo, H., Fujii, C.Y., Utterback, T.R., Barnstead, M.E., Bowman, C.L., White, O., Nierman, W.C. and Fraser, C.M.				
TITLE	Arabidopsis thaliana chromosome III BAC F16B3 genomic sequence				
JOURNAL	Unpublished				
MEDLINE					
PUBMED					
REFERENCE	2 (bases 1 to 103904)				
AUTHORS	Lin, X. and Kaul, S.				
TITLE	Direct Submission				
JOURNAL	Submitted (18-JAN-2000) The Institute for Genomic Research, 9712 Medical Center Dr., Rockville, MD 20850, USA, xlin@tigr.org				
COMMENT	3 (bases 1 to 103904)				
	Lin, X.				
	Direct Submission				
	Submitted (24-JAN-2001) The Institute for Genomic Research, 9712 Medical Center Dr., Rockville, MD 20850, USA				
	On Jan 24, 2001 this sequence version replaced gi:12280770.				
	Address all correspondence to: at@tigr.org				
	BAC clone F16B3 is from Arabidopsis chromosome III and is near the molecular marker mi74.				
	The orientation of the sequence is from SP6 to T7 end of the BAC clone.				
	Genes were identified by a combination of three methods: Gene prediction programs including GRAIL (available by anonymous ftp from arthur.epm.ornl.gov), GeneFinder (Phil Green, University of Washington), Genscan (Chris Burge, http://gnomic.stanford.edu/~chris/GENSCANW.html), and NetPlantGene (http://www.cbs.dtu.dk/netgene/cbsnetgene.html), searches of the complete sequence against a peptide database and the Arabidopsis EST database at TIGR (http://www.tigr.org/tdb/at/at.html).				
	Annotated genes are named to indicate the level of evidence for their annotation. Genes with similarity to other proteins are named after the database hits. Genes without significant peptide similarity but with EST similarity are named as 'unknown' proteins. Genes without protein or EST similarity, that are predicted by more than two gene prediction programs over most of their length are annotated as 'hypothetical' proteins. Genes encoding tRNAs are predicted by tRNAscan-SE (Sean Eddy, http://genome.wustl.edu/eddy/tRNAscan-SE/). Simple repeats are identified by repeatmasker (Arian Smit, http://ftp.genome.washington.edu/RM/RepeatMasker.html).				
FEATURES	Location/Qualifiers				
	1..103904				
	/organism="Arabidopsis thaliana"				
	/cultivar="Columbia"				
	/db_xref="taxon:3702"				
	/chromosome="III"				
	/map="mi74"				
	/clone="F16B3"				
	<129..>979				
	/gene="F16B3.1"				
	/note="predicted by genefinder"				
	join(<129..241,586..>979)				
	/gene="F16B3.1"				
	join(129..241,586..979)				
	/gene="F16B3.1"				
	/note="hypothetical protein"				
	/codon_start=1				
	/protein_id="AAF32445.1"				
	/db_xref="GI:6957701"				
	/translation="MEKDWGSGPLSENEGALLCYDDEYQFGGAPAKLPQRYLSEIG				
	ELQLLGDDDDVYSLKDYGEIAGYKCCNVEYRYLGLVILGRHGVPTTKY				
	AVNTTPSDESLCAAEEFODRDSVTKLLSDMHICDARPVDFVYLPNSQFKSSPGEP				
	SFVTCFSG"				
	complement(<3079..>4265)				
	/gene="F16B3.2"				
	/note="identical to COL2 GB: AAB67879 from [Arabidopsis thaliana]"				
	complement(join(<3079..3438,3582..>4265))				
	/gene="F16B3.2"				

```

CDS
complement(join(3079..3438,3582..4265))
/gene="F16B3.2"
/codon_start=1
/product="COL2"
/protein_id="AAF32446.1"
/db_xref="GI:6957702"

repeat_region
/translation="MKESNESGTWARACDTCRSAACTVYCEADSAYLCTTCDAHVH
AANVRARRRVVSCSAPAFCLKADAAASLCTACDAEIHSAANPLARRHQRVPI
PLSANSCEMAPSETDAONEDREVASWLLPGKNGIGNQNGNGLFGVEYLDLVDS
SSMDNQFEDQYTHVQRSGGQVVPLOVEESTSHLOQSQONFOLGNGFGSSGAHYN
NNSLDLHNSASVSMDSVPESTASDITVQHPRTKETIDLSGPTQVVOQLTPM
ERARVLRREKKTKRFDKTRIARSAKAIERPRKGRFAKRIETAEAEIEFSTS
LMSYGYGIVPSP"
5228..5255

repeat_region
/translation="MELPATKLRIGLVIPLILLITAPILYLFYFGVPLYYSTYKHL
SNSSSSAISPSRNYNHHSSSDSKYKTESSESSDNDYDDYTHDKSSSLNDRSLS
ISSNGHHQVTPKKEHRRKKRRKCDIFSGEWIPNPKADYITNTTCRAIHEHQCILY
EHPEDSQOEFNFQWKYNTFTATFTVHLVRAEETETGTPGNSFYNYLDEPD
PWAQSIGEDFYIISSQGWFFRPLFLDKQKRGICLYCIPGVNRVGAFAFYRRLR
TFKFTILGLENFKGEVLFRTAPSHFEGGEWKGAMRQS"
complement(17673..17778)
/repeat_region
/repeat_family="(CA)n"
complement(18048..18074)
/repeat_family="(TA)n"

Query Match 100.0%; Score 180; DB 8; Length 103904;
Best Local Similarity 100.0%; Pred No. 1.8e-44;
Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCGCGGGAATAATCTTATCTGATGCGAGTTGTTGACGAACCTCTTTTACAACCGA 60
Db 62689 ATGGCGCGGGAATAATCTTATCTGATGCGAGTTGTTGACGAACCTCTTTTACAACCGA 62748

QY 61 ATCGTTCTGAGTCATCTTTTGGCGGGAATCTATGGAACCTTACCTATTTCTCCAG 120
Db 62749 ATCGTTCTGAGTCATCTTTTGGCGGGAATCTATGGAACCTTACCTATTTCTCCAG 62808

QY 121 ACATGCTCCGAAATTAACCTCGCCGGAACCTACTACTTCTATCTCCGTTTCTCTGG 180
Db 62809 ACATGCTCCGAAATTAACCTCGCCGGAACCTACTACTTCTATCTCCGTTTCTCTGG 62868

RESULT 5
ATRNASC5D
LOCUS A.thaliana mRNA for sterol-C5-desaturase.
DEFINITION X90454
ACCESSION X90454
VERSION X90454.1
KEYWORDS sterol-C5-desaturase gene.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

```


Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 1175)
Gachotte, D., Hüsselstein, T., Bard, M., Lacroix, F., and Benveniste, P.
Isolation and characterization of an Arabidopsis thaliana cDNA
encoding a delta 7-sterol-C-5-desaturase by functional
complementation of a defective yeast mutant
Plant J. 9 (3), 391-398 (1996)
97077348
PUBMED
8919915
REFERENCE
2 (bases 1 to 1175)
Gachotte, D.
Direct Submission
Submitted (12-JUL-1995) D. Gachotte, Institut de Biologie
Moléculaire, Enzymologie Moléculaire et Cellulaire,
Institut de botanique 28, rue Goethe, F- 67083 Strasbourg, FRANCE
JOURNAL
FEATURES
source
1. .1175
/organism="Arabidopsis thaliana"
/variety="Landsberg erecta"
/db_xref="taxon:3702"
118..963
/codon_start=1
/product="sterol-C5-desaturase"
/protein_id="CAA62079.1"
/db_xref="GI:1061038"
/db_xref="SPTREMBL:O39208"
/translation="MAADNAYLMQFVDETSEFYRNLVLSHLLPANLWELPHFLQTLWR
NYLGLTLLIISFLWCFYIYRKNILVLPKDALPTIKAMRQFQVAMKAMPWTLLP
TVSEMERQWTCFASIGFQWLYEYVIAIYLVFVEFGIYWMHRLHDKPLRYKL
HATHYIKQNTLSPPFAGLPHFDGLQAVPHVIAIYVIFHTTIGLLFWEAIA
ANTHDCIHRGNIWPMVGAGYHTIHTYKHYGHTIWMWFMGSLRDLPLEEDNDKDS
FKAE"
BASE COUNT 300 265 c 227 g 383 t
ORIGIN
Query Match 99.1%; Score 178.4; DB 8; Length 1175;
Best Local Similarity 99.4%; Pred. No. 5.1e-44;
Matches 179; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Oy 1 ATGGCGCGGATATGCTTATCTGTCAGTGTGTTGACGAAACCTCTTTTACACCGA 60
|||||
Db 118 ATGGCGCGGATATGCTTATCTGTCAGTGTGTTGACGAAACCTCTTTTACACCGA 177
Oy 61 ATCGCTTCTGACTATCTTTTGGCGGCAATCTATGGAACCTTACCTCATTTTCTCCAG 120
|||||
Db 178 ATCGCTTCTGACTATCTTTTGGCGGCAATCTATGGAACCTTACCTCATTTTCTCCAG 237
Oy 121 ACATGGTCCGAAATACCTGCGCGGAACCTACTATCTATCTCCGGTTTCTCTGG 180
|||||
Db 238 ACATGGTCCGAAATACCTGCGCGGAACCTACTATCTATCTCCGGTTTCTCTGG 297
RESULT 6
AF081794
LOCUS
DEFINITION
AF081794
ACCESSION
AF081794.1 GI:4140397
VERSION
KEYWORDS
SOURCE
Nicotiana tabacum.
Nicotiana tabacum
Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.
1 (bases 1 to 1155)
Hüsselstein, T., Schaller, H., Gachotte, D., and Benveniste, P.
Delta7-sterol-C-5-desaturase: molecular characterization and
functional expression of wild-type and mutant alleles
Plant Mol. Biol. 39 (5), 891-906 (1999)
99273992
PUBMED
10344195
REFERENCE
2 (bases 1 to 1155)

Benveniste, P.
Direct Submission
Submitted (30-JUL-1998) Plant Molecular Biology Institute, Centre
National de la Recherche Scientifique, 28 rue Goethe, Strasbourg
67084, France
JOURNAL
FEATURES
source
1. .1155
/organism="Nicotiana tabacum"
/cultivar="xanthi Sh6"
/db_xref="taxon:4097"
/tissue_type="calli derived from leaf protoplasts"
103..918
/codon_start=1
/product="sterol-C5(6)-desaturase"
/protein_id="AAD04034.1"
/db_xref="GI:4140398"
/translation="MEDYILKRFVETSEFYRNLVLTGTPMPSWNGPLPHMLQGLRNLY
GGVLLYFISGLWCFYIYKRNLYIYPRDAIPSNKMLLOISVAMKAMPWTCALPSS
EYIENGWTKCFARISDVGLSYVIAIYLVFVEFGIYWMHRLHDKPLRYKHAT
HHYIKQNTLSPPFAGLPHFDGLQAVPHVIAIYVIFHTTIGLLFWEAIA
HDCIHRGNIWPMVGAGYHTIHTYKHYGHTIWMWFMGSLRDLPLEEDNDKDS
FKAE"
BASE COUNT 310 a 239 c 245 g 361 t
ORIGIN
Query Match 45.6%; Score 82; DB 8; Length 1155;
Best Local Similarity 66.3%; Pred. No. 1.9e-14;
Matches 118; Conservative 0; Mismatches 60; Indels 0; Gaps 0;
Oy 3 GCGCGGATATGCTTATCTGTCAGTGTGTTGACGAAACCTCTTTTACACCGAAT 62
|||||
Db 96 GTCGAGATGAGGATTACTTGAACACATTTGTGAGGAGACATCGTTTACAATCGGTT 155
Oy 63 GGTCTGAGTCATCTTTTGGCGGCAATCTATGGAACCTTACCTCATTTTCTCCACAC 122
|||||
Db 156 GGTCTGAGTCATCTTTTGGCGGCAATCTATGGAACCTTACCTCATTTTCTCCACAC 215
Oy 123 ATGGCTCCGAAATACCTGCGCGGAACCTACTATCTATCTCCGTTTCTCTCG 180
|||||
Db 216 ATGGCTCCGAAATACCTGCGCGGAACCTACTATCTATCTCCGTTTCTCTCG 273
RESULT 7
AF003214
LOCUS
DEFINITION
Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 1,
BAC clone:OSJNBa0083M16.
AP003214
ACCESSION
AP003214.3 GI:15528580
VERSION
KEYWORDS
SOURCE
Oryza sativa (japonica cultivar-group) (cultivar:Nipponbare) DNA,
clone:OSJNBa0083M16.
Oryza sativa (japonica cultivar-group)
Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
1
Sasaki, T., Matsumoto, T. and Yamamoto, K.
Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, BAC
clone:OSJNBa0083M16
Published Only in database (2001)
2 (bases 1 to 179428)
Sasaki, T., Matsumoto, T. and Yamamoto, K.
Direct Submission
Submitted (19-FEB-2001) Takuji Sasaki, National Institute of
Agrobiological Sciences, Rice Genome Research Program; Kannondai
2-1-2, Tsukuba, Ibaraki 305-8602, Japan
(E-mail:tsasaki@nias.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/,
Tel:81-298-38-7441, Fax:81-298-38-7468)
On Sep 7, 2001 this sequence version replaced gi:13620983.
Genes were predicted from the integrated results of the following:
GENSCAN1.0, BLASTN2.0, BLASTX2.0 as well as SplicePredictor
(October 1998 version). The genomic sequence was searched against
NCBI Nonredundant Protein database, nr

(ftp://ncbi.nlm.nih.gov/blast/db) and the cDNA sequence database at RGP. Protein homologies of the coding regions were searched against NCBI NonRedundant Protein database with BLASTp2.0. ESTs represent the identified cDNA sequences using BLASTN 2.0 with the corresponding DDBJ accession no. and RGP clone ID.

A gene with identity or significant homology to a protein is classified based on the protein name to indicate the homology level such as same name, 'putative-' and '-like protein'. A gene without significant homology to any protein but with EST homology (covering almost the entire length of partial sequence) is classified as an 'unknown' protein. A gene predicted with a gene prediction program is classified as a 'hypothetical' protein.

The orientation of the sequence is from -21M13 to M13rev of the BAC clone. This sequence of OSJNBa0083M16 clone has an overlap with P504H10 (DDBJ: AP002526) clone at the position 1 to 42,554 and with P0443D08 (DDBJ: AP003250) at the position 178,805 to 179,428 of 3' end. The sequence of this clone starts at the position 100,962 of P504H10 and ends at the position 624 of P0443D08. Detailed information on overlap and assembly quality together with annotation of this entry is available at <http://rgp.dna.affrc.go.jp/GenomesSeq.html>.

FEATURES

Source

```
1..179428
/organism="Oryza sativa (japonica cultivar-group)"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/chromosome="1"
/clone="OSJNBa0083M16"
/complement(join(870..1087,1280..1364))
/gene="OSJNBa0083M16.1"
/complement(join(870..1087,1280..1364))
/gene="OSJNBa0083M16.1"
/notes="contains EST AU092803(C53725)"
/codon_start=1
/product="similar to Zea mays Bowman-Birk proteinase
inhibitors WIPI precursor"
/protein_id="BAB64603.1"
/db_xref="GI:15528581"
/translation="MKSSTLLVLAQAALINGIFAAVAKENAVGESKAIDNPGLK
CCSNCNFEFSGLYTCDIVKCDPVCCKAVVTKYPMKFKCTDFLGMGPPCKH"
5689..9051
/gene="OSJNBa0083M16.2"
5689..9051
/gene="OSJNBa0083M16.2"
/codon_start=1
/product="similar to Lycopersicon pimpinellifolium Cf-2
gene"
/protein_id="BAB64604.1"
/db_xref="GI:15528582"
/translation="MYSDFEPALFSLTSLQRLDLSMNSLGTSTTKDAEFDRLTSLT
HNLNSGLDGGQIPMGINKILNLSVLSKRYVNDNSDISFNESDDEIIFTGDSYNHL
QESRLMSLVENSLEKLYLDHVDNSTVDDCKTLAQSVPLQVLSLDGCSLNTPIH
HSLLRLSLTILNLSQSGIAVNLFPDEPMGANLTVLRSNNLEGGWPPDKFELKLN
LRILDSLNNMLGLHLPKVTSLTLRLGTFSTAKRISSEFNMLKELGLEGLIS
KDFLSFGLIWSLCHLELNSLGLSDLSLWSIAGKNTLLISLSEFSTKPS
ISNFKRLSLWFLGCLNLRPIMSAIGDLVLSLQSLNSCNTYSNPSIGNITNKL
YNSPGFGLPMPAALGNLKSWSFNCEFTGMPSTIGNTLQTLLEIAACRFSGP
IYSGOLKELRALFEGCNMSGRIPNSVTNNMSKLYLGLPANYLSGKTPARLFTLPA
LLFLDLFGNHFSGPIQEDAVPSYLSLQSLTSLNETGPEPFFELTSLIALEIDLNN
LAGSVDSLFRKLRLDLNLSHNNLSVIMDEGNSNTLSYSELKELGACCNITKF
PILRLSDMSYLDLSCNKISGNIKPWIKWSSVSHNLNSHNLMTSMEVASYLLPF
NRHFTLDSLNSLQAGQIPNLNSAFDLSHNAFSSILPNFTLYLSKTWLSMKN
ISGNPISCTNSLSLVLAHNNFSGPPSCIMEQTYFNILNLRGNHPEGLMPTNVT
RCAFT IDLNGK IEGRPLRAGNCTYLEVLDLGNKADTFPSMLGSLSNLRLVLR
SNRLVSGTYETEDSGDHPNLIIDLASNNFTSLGHPQEFKISMKYNNGTGETI
SHRSISDGYDTVTISCKFSMTFRLITLITLIDLSNDALEGSIPESVGLKLSLH
VLNLSHNAFSGPIQIGITALELSDLSNNWISGEIPELTNLTFLTLNLSNNOLE
KNLPSRFQFATPENSIEGNAGLGDPLPKCASWSPSPSAEPHVSSESHVDITVMFLFY
GVGFGVAGVGLMKTSLWINRWFHSAVSQRQT"
/complement(join(9685..9961,10319..10662))
/gene="OSJNBa0083M16.3"
/complement(join(9685..9961,10319..10662))
/gene="OSJNBa0083M16.3"
```

```
/note="hypothetical protein"
/codon_start=1
/protein_id="BAB64605.1"
/db_xref="GI:15528583"
/translation="MESPRSCDHAAISLVWGTRRRLLHIRMAELAAEPGSPQDLSL
QLSMRQHLGCADDKADATACGIELHGLAWPRTPPCAATTCAARRCASPCGAAR
SSAARAMAAGVCRNVSVPYLRPGAAVNNPPLAPASAPSPVRRPRRRPSVAT
ISARGGGEERGAAYVEGGRRAPAAVEGVPAPALGAVAAWCGGGG"
/complement(join(12004..12746,13413..13449))
/gene="OSJNBa0083M16.4"
/complement(join(12004..12746,13413..13449))
/gene="OSJNBa0083M16.4"
/notes="hypothetical protein"
/codon_start=1
/protein_id="BAB64606.1"
/db_xref="GI:15528584"
/translation="MTVLGRKOSGLSVAVAPMEELAPELIGPRISFSHDGVVATAGA
ATPAARSDTLVVMSSWRRLPEPEDFANAAADVAPADRLFAAGKLLPVPLPPVH
PKPSPCKQQQATSGGGVKPTSVQRPQSWTSPTTRSCSVNSATTAPRSGSGSFCPS
FPLMRSRAGSAAAGGGLGVGVSGGHHRPQHKKAGATAAAYYGGSRNGSGHG
VRVSPVINVPISGTSVMNMLSYLLDCGCKTKNRGFGLNC"
/complement(join(16789..17270,17559..17837,17983..18106,
18193..18498,18750..18902))
/gene="OSJNBa0083M16.5"
/complement(join(16789..17270,17559..17837,17983..18106,
18193..18498,18750..18902))
/gene="OSJNBa0083M16.5"
/notes="hypothetical protein"
/codon_start=1
/protein_id="BAB64607.1"
/db_xref="GI:15528585"
/translation="MGRLLLVSLDRRRRYIVLPSCQGVDPVVCGRNAQRFQGH
DYHAGFERPKPLLSGLSPRFRVPLRHPHIAKNSILMEEGKHAIVDLSEBERRG
CGAPGEGSDEVDGDAVLGRGAESTPEGSADVDDAGHGVAAEKANDDLEPLEPE
GYTVAGSGGGDGTASDGGENELHEVPKNPMATREAFRATDANNTVSTYRGLMSRKK
NAGPTRFVSAGAPPSPALSTPALSLSSELKSSPENASVADPADHGASPGRNGDOL
EAAKQSEARFAPDAAAAAAASHREPSKQKQPRPERFIEPEGEAAARAKACRS
GIALDRFTISQLNPSGSTEWEVETADVGGGEGECTSDPSCSIAISDSGPP
EPLDPDRRIYSVLAVLGVSLAVSMVALFVIFGSESPSPSPNQVIFKQY"
join(21657..22226,22293..22395,23106..23221,24403..24457,
25399..25589)
/gene="OSJNBa0083M16.6"
join(21657..22226,22293..22395,23106..23221,24403..24457,
25399..25589)
/gene="OSJNBa0083M16.6"
/notes="hypothetical protein"
/codon_start=1
/protein_id="BAB64608.1"
/db_xref="GI:15528586"
/translation="MVISAEHPHPPPPDPVADCAAAVEAAASARHQAALVSFPFLK
TWGSHRVLCRCAHVNAGDAIAAARRSPQLDGVKERLLHLREVDSGVVADAEEVAA
VAAAPAAAASRPWKLRSRRRPAVAPASSAPPERRAARAEALDRARFVTLTSE
ETEDVAVTGARPRRPRRPVKQLDGPTRNFVKTHVPSCYFPVHGCRSPSPSP
TCGPTIVPRGTHLSEELNWDYTLTGCGRGRGQVCSASRTKKGKTHNNNEHSF
LSPWNYLEPGEVELDALRCQLSCSDCAPRPSRSSRSIYINQHLQAMFFLHGQD
RLGGTGICND"
/complement(join(30012..30245,31517..31582,31671..31817))
/gene="OSJNBa0083M16.7"
/complement(join(30012..30245,31517..31582,31671..31817))
/gene="OSJNBa0083M16.7"
/notes="contains ESTs AU075676(E106083).AU075677(E106083)"
/codon_start=1
/product="similar to Mus musculus SURF-5 gene"
/protein_id="BAB64609.1"
/db_xref="GI:15528587"
/translation="MSKAGAAAGTAATAAAAAVOKTKLLKADADVSSLYDNFAAF
INIARVNPPVNTQEAQMDMRGRLHSADSLKLSLSELRKTAIFISGLASUTENV
RIETLSQAQAGTQERMLRIGEGAAAGSLKEALHYSSVYRTPPDE"
/complement(42717..43520)
/gene="OSJNBa0083M16.8"
/complement(42717..43520)
/gene="OSJNBa0083M16.8"
/notes="hypothetical protein"
```

[illegible]

TITLE
JOURNAL

Direct Submission
Submitted (20-JAN-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 174556)
Worley, K.C.

REFERENCE
AUTHORS
TITLE
JOURNAL

Direct Submission
Submitted (13-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Jul 12, 2002 this sequence version replaced gi:18846184.

COMMENT

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GIVE
Center clone name: CH230-103L10
----- Summary Statistics
Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 114370 bases at least Q40
Consensus quality: 121991 bases at least Q30
Consensus quality: 127877 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 71 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence.
* as soon as it is available and the accession number will
* be preserved.

1 1241: contig of 1241 bp in length
1242 1341: gap of unknown length
1342 2387: contig of 1046 bp in length
2388 2487: gap of unknown length
2488 3645: contig of 1158 bp in length
3646 3745: gap of unknown length
3746 5151: contig of 1406 bp in length
5152 5251: gap of unknown length
5252 6629: contig of 1378 bp in length
6630 6729: gap of unknown length
6730 8303: contig of 1574 bp in length
8304 8403: gap of unknown length
8404 9650: contig of 1247 bp in length
9651 9750: gap of unknown length
9751 10767: contig of 1017 bp in length
10768 10867: gap of unknown length
10868 12522: contig of 1655 bp in length
12523 12622: gap of unknown length
12623 14174: contig of 1552 bp in length
14175 14274: gap of unknown length
14275 15962: contig of 1688 bp in length
15963 16062: gap of unknown length
16063 17144: contig of 1082 bp in length
17145 17444: gap of unknown length
17446 18363: contig of 1119 bp in length
18364 18463: gap of unknown length
18464 19715: contig of 1252 bp in length
19716 19815: gap of unknown length
19816 21377: contig of 1562 bp in length
21378 21477: gap of unknown length
21478 22694: contig of 1217 bp in length
22695 22794: gap of unknown length
22795 24119: contig of 1325 bp in length
24120 24219: gap of unknown length
24220 25302: contig of 1083 bp in length
25303 25402: gap of unknown length

25403 26782: contig of 1380 bp in length
26783 26882: gap of unknown length
26883 28936: contig of 2054 bp in length
28937 29036: gap of unknown length
29037 30973: contig of 1837 bp in length
30974 32115: gap of unknown length
32116 32415: contig of 1442 bp in length
32416 34135: gap of unknown length
34136 34235: gap of unknown length
34236 35785: contig of 1550 bp in length
35786 35885: gap of unknown length
35886 37119: contig of 1233 bp in length
37119 37219: gap of unknown length
37219 38557: contig of 1639 bp in length
38558 38957: gap of unknown length
38957 40822: contig of 1665 bp in length
40822 40922: gap of unknown length
40922 42897: contig of 2175 bp in length
42897 42998: gap of unknown length
42998 44334: contig of 1436 bp in length
44334 44533: gap of unknown length
44533 46874: contig of 2341 bp in length
46874 46974: gap of unknown length
46974 48922: contig of 1948 bp in length
48922 48923: gap of unknown length
48923 51119: contig of 2097 bp in length
51119 51220: gap of unknown length
51220 53055: contig of 1836 bp in length
53055 53156: gap of unknown length
53156 55184: contig of 2029 bp in length
55184 55284: gap of unknown length
55284 56390: contig of 1106 bp in length
56390 56490: gap of unknown length
56490 58207: contig of 1717 bp in length
58207 58307: gap of unknown length
58307 61073: contig of 2766 bp in length
61073 61174: gap of unknown length
61174 62294: contig of 1121 bp in length
62294 62394: gap of unknown length
62394 63933: contig of 1539 bp in length
63933 64033: gap of unknown length
64033 66199: contig of 2166 bp in length
66199 66299: gap of unknown length
66299 68544: contig of 2245 bp in length
68544 68644: gap of unknown length
68644 71314: contig of 2670 bp in length
71314 71414: gap of unknown length
71414 73505: contig of 2091 bp in length
73505 73606: gap of unknown length
73606 76323: contig of 2718 bp in length
76323 76423: gap of unknown length
76423 78745: contig of 2322 bp in length
78745 78845: gap of unknown length
78845 80957: contig of 2112 bp in length
80957 82339: contig of 1282 bp in length
82339 82439: gap of unknown length
82439 84739: contig of 2300 bp in length
84739 84839: gap of unknown length
84839 87042: contig of 2203 bp in length
87042 87142: gap of unknown length
87142 90257: contig of 3115 bp in length
90257 90357: gap of unknown length
90357 92861: contig of 2504 bp in length
92861 92961: gap of unknown length
92961 95075: contig of 2114 bp in length
95075 95175: gap of unknown length
95175 99198: contig of 4023 bp in length
99198 99298: gap of unknown length
99298 102680: contig of 3382 bp in length
102680 102680: contig of 3382 bp in length

Query Match

19.9%; Score 35.8; DB 2; Length 174556;

```

Best Local Similarity 53.1%; Pred. No. 3;
Matches 76; Conservative 0; Mismatches 67; Indels 0; Gaps 0;

QY 38 ACAGAACTCTCTTTTACACACGAATCGTTCTGAGTCATCTTTGCGGCGAATCATGCG 97
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 36620 ACAATGACCTCTTTTACATATATATTCATTCATGCACATCATGCTGATGATACCTATCTC 36679

QY 98 AACCCCTACCTGATTTTCCAGACATGCTCCGAAATACCTCCCGGAACCCACCTACAT 157
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 36680 AACACTTACCAATCTCTCCATCTTACTCCCGCAGAACCTTCTCTCCCAATGGA 36739

QY 158 ACTTCATCTCCGTTTCTCTCTGG 180
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 36740 AACACTGTTTCTGCTGCTTTGG 36762

RESULT 12
AC007247/c
LOCUS Homo sapiens BAC clone RP11-305H21 from Y, complete sequence.
DEFINITION AC007247
ACCESSION AC007247
VERSION AC007247.5 GI:10801460
KEYWORDS HTG.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 116788)
Sulston, J.E. and Waterston, R.
Toward a complete human genome sequence
Genome Res. 8 (11), 1097-1108 (1998)
99063792
9847074
2 (bases 1 to 116788)
Kalicki, J., Bemis, G. and Hawrysko, C.
The sequence of Homo sapiens BAC clone RP11-305H21
Unpublished
3 (bases 1 to 116788)
Waterston, R.H.
Direct Submission
Submitted (07-APR-1999) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
4 (bases 1 to 116788)
Waterston, R.
Direct Submission
Submitted (14-OCT-2000) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Oct 14, 2000 this sequence version replaced gi:9838137.
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc
Contact: sapiens@wustl.wustl.edu
----- Summary Statistics
Center project name: H_NH0305H21
-----

```

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

The position of this clone was established as part of a

collaboration between the Human Chromosome Y Mapping Project (Tomoko Kawaguchi, Helen Skaletsky, Laura G. Brown, Steve Rozen, and David C. Page at the Whitehead Institute for Biomedical Research, Cambridge MA) and the Washington University Genome Sequencing Center, St. Louis MO.

SOURCE INFORMATION:

The RP11-11 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Moon, P.Y., Zhao, B., Frengen, E., Taten, M., Catanese, J.J. and de Jong, P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at the Roswell Park Cancer Center Institute (<http://bacpac.med.buffalo.edu>)

VECTOR: pBACe3.6

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is RP11-558K21, 200 bp overlap; the clone sequenced to the right is RP11-105L10, 200 bp overlap. Actual start of this clone is at base position 156866 of RP11-558K21; actual end is at base position 53998 of RP11-105L10.

The sequence RP11-305H21 from base position 114735 to 116434 contains a tandem repeat. The region was sized with PCR from clone DNA.

FEATURES

```

source
1..116788
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="Y"
/clone="RP11-305H21"
/clone_lib="RPCI-11"
/repeat_region
3076..4236
/rpt_family="L1"
/repeat_region
4244..4318
/rpt_family="L1"
/repeat_region
4506..4807
/rpt_family="Alu"
/repeat_region
5517..5697
/rpt_family="Alu"
/repeat_region
5916..6328
/rpt_family="L1"
/repeat_region
6602..6677
/rpt_family="MER1_type"
/repeat_region
6714..6806
/rpt_family="MER1_type"
/repeat_region
6834..6986
/rpt_family="MER1_type"
/repeat_region
7529..7935
/rpt_family="MaLR"
/repeat_region
7982..8316
/rpt_family="L1"
/repeat_region
8995..9364
/rpt_family="L1"
/repeat_region
9365..9654
/rpt_family="Alu"
/repeat_region
9655..9670
/rpt_family="L1"
/repeat_region
9666..9846
/rpt_family="L1"
/repeat_region
10008..10584
/rpt_family="ERV1"
/repeat_region
10600..10717
/rpt_family="L1"
/repeat_region
10793..10950
/rpt_family="ERV1"
/repeat_region
12680..13210
/rpt_family="MER1_type"
/repeat_region
13241..13466
/rpt_family="MaLR"
/repeat_region
14411..14515
/rpt_family="Alu"

```


Query Match 19.3%; Score 34.8; DB 9; Length 116788;
Best Local Similarity 51.3%; Pred. No. 6.1;
Matches 81; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

Qy 18 TTATCTGATCGAGTTTGTTCCAGCAAAACCTCTTTTTTACAACCAGTAATGGTTCTGACTCATCT 77
|| | ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 50771 TTCTTCAGACATTTCTCCTGAACCAACAGTAGTATTCAATCTGCCCTACTAGTGCATCA 50711

Qy 78 TTTGCGGGCGAATCTATGGGAACCCCTTACCTCATTTTTTCTCCAGCACATGGCTCCGAAATTA 137
|| | ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 50711 TTAGCGAGTTGTGTTATTTCCTCTATTAACTGGGTTTTTCCCAAAGAAGAGGACCTTG 50651

138 CCTGCGGGGAACCTACTATACTTCTATCTCCGGTTTCCC 175
| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 50651 CTAGCAAAAGTCCTTTCCAGAAGTCACATCCATTTTAC 50614

RESULT 13
AC116209
LOCUS
DEFINITION Rattus norvegicus clone CH230-16F3, *** SEQUENCING IN PROGRESS ***, linear HTG 13-JUL-2002
61 unordered pieces.
AC116209
VERSION AC116209.2 GI:21736254
KEYWORDS HTG; HTGS_PHASE1.
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE 1 (bases 1 to 186537)
AUTHORS Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C., Alsbrooks,S.L., Amarantunge,H.C., Are,J.R., Ayele,M., Banks,T., Barbara,J., Benton,J., Bimaga,K., Blankenburg,K., Bonnin,D., Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Bunay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Caron,T.F., Carter,M., Cavazos,S.R., Chacko,J.J., Chavez,D., Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B., Homsí,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L., Li,Z., Licharge,O., Lieu,C., Liu,J., Liu,W., Loulseghe,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E., Massey,E., Mawhinney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokwenkwos,S., Ogum,M., Okwuonu,G., Oraguanye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L., Pickens,R., Primeau,E., Pu,L., Quiles,M., Ren,Y., Rivers,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savory,G., Scherer,S., Scott,G., Shen,H., Shoohtari,N., Sisson,I., Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svatek,A., Taborp,T., Tamerisa,A., Tamerisa,K., Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q., Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S., Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D., Weinstein,G., and Gibbs,R.

TITLE Direct Submission
JOURNAL Unpublished

Direct Submission
Unpublished

REFERENCE
AUTHORSTITLE
JOURNAL

2 (bases 1 to 186537)

Worley, K.C.

Direct Submission

Submitted (26-MAR-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 186537)

Worley, K.C.

Direct Submission

Submitted (13-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

On Jul 12, 2002 this sequence version replaced gi:19718403.

COMMENT

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GWK

Center clone name: CH230-16F3

----- Summary Statistics

Sequencing vector: Plasmid;

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.990329

Consensus quality: 143677 bases at least Q40

Consensus quality: 148892 bases at least Q30

Consensus quality: 153855 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length

* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).

* NOTE: This is a 'working draft' sequence. It currently

* consists of 61 contigs. The true order of the pieces

* is not known and their order in this sequence record is

* arbitrary. Gaps between the contigs are represented as

* runs of N, but the exact sizes of the gaps are unknown.

* This record will be updated with the finished sequence

* as soon as it is available and the accession number will

* be preserved.

* 1 1503: contig of 1503 bp in length

* 1504 1603: gap of unknown length

* 1604 2662: contig of 1059 bp in length

* 2663 2762: gap of unknown length

* 2763 4053: contig of 1291 bp in length

* 4054 4153: gap of unknown length

* 4154 5498: contig of 1345 bp in length

* 5499 5598: gap of unknown length

* 5599 6869: contig of 1271 bp in length

* 6870 6969: gap of unknown length

* 6970 8142: contig of 1173 bp in length

* 8143 8242: gap of unknown length

* 8243 9430: contig of 1188 bp in length

* 9431 9530: gap of unknown length

* 9531 10878: contig of 1348 bp in length

* 10879 10978: gap of unknown length

* 10979 12038: contig of 1060 bp in length

* 12039 12138: gap of unknown length

* 12139 13914: contig of 1776 bp in length

* 13915 14014: gap of unknown length

* 14015 15369: contig of 1355 bp in length

* 15370 15469: gap of unknown length

* 15470 16954: contig of 1485 bp in length

* 16955 17054: gap of unknown length

* 17055 18245: contig of 1191 bp in length

* 18246 18345: gap of unknown length

* 18346 20838: contig of 2493 bp in length

* 20839 22434: gap of unknown length

* 22435 22534: contig of 1496 bp in length

* 22535 24082: gap of unknown length

* 24083 24182: contig of 1548 bp in length

* 24183 25867: gap of unknown length

* 25868 25967: gap of unknown length

* 25968 28126: contig of 2159 bp in length

* 28127 28226: gap of unknown length

* 28227 29816: contig of 1590 bp in length

* 29817 29917: gap of unknown length

* 29918 32638: contig of 2722 bp in length

* 32639 32738: gap of unknown length

* 32739 35012: contig of 2274 bp in length

* 35013 35112: gap of unknown length

* 35113 36465: contig of 1353 bp in length

* 36466 36565: gap of unknown length

* 36566 39381: contig of 2816 bp in length

* 39382 39481: gap of unknown length

* 39482 41964: contig of 2483 bp in length

* 41965 42064: gap of unknown length

* 42065 43709: contig of 1645 bp in length

* 43710 43809: gap of unknown length

* 43810 46620: contig of 2811 bp in length

* 46621 46720: gap of unknown length

* 46721 48478: contig of 1758 bp in length

* 48479 48578: gap of unknown length

* 48579 51044: contig of 2466 bp in length

* 51045 51144: gap of unknown length

* 51145 52706: contig of 1562 bp in length

* 52707 52806: gap of unknown length

* 52807 55233: contig of 2427 bp in length

* 55234 55333: gap of unknown length

* 55334 57119: contig of 1786 bp in length

* 57120 57219: gap of unknown length

* 57220 59901: contig of 2682 bp in length

* 59902 60001: gap of unknown length

* 60002 62113: contig of 2112 bp in length

* 62114 62213: gap of unknown length

* 62215 64247: contig of 2034 bp in length

* 64248 64347: gap of unknown length

* 64348 66015: contig of 1668 bp in length

* 66016 66115: gap of unknown length

* 66116 69403: contig of 3288 bp in length

* 69404 69503: gap of unknown length

* 69504 72038: contig of 2535 bp in length

* 72039 72138: gap of unknown length

* 72139 75834: contig of 3696 bp in length

* 75835 75934: gap of unknown length

* 75935 79004: contig of 3070 bp in length

* 79005 79104: gap of unknown length

* 79105 82240: contig of 3136 bp in length

* 82241 84599: contig of 2259 bp in length

* 84600 84699: gap of unknown length

* 84700 88192: contig of 3493 bp in length

* 88193 88292: gap of unknown length

* 88293 91473: contig of 3181 bp in length

* 91474 91573: gap of unknown length

* 91574 94266: contig of 2693 bp in length

* 94267 94366: gap of unknown length

* 94367 98103: contig of 3737 bp in length

* 98104 98203: gap of unknown length

* 98204 101269: contig of 3066 bp in length

* 101270 101369: gap of unknown length

* 101370 104746: contig of 3377 bp in length

* 104747 104846: gap of unknown length

* 104847 108037: contig of 3191 bp in length

* 108038 108137: gap of unknown length

* 108138 111967: contig of 3830 bp in length

* 111968 112067: gap of unknown length

* 112068 115144: contig of 3077 bp in length

* 115145 115245: gap of unknown length

* 115246 117937: contig of 2693 bp in length

* 117938 118037: gap of unknown length

* 118038 121584: contig of 3547 bp in length

* 121585 121685: gap of unknown length

* 121686 127262: contig of 5578 bp in length

* 127263 131103: gap of unknown length

* 131104 131103: contig of 3741 bp in length

Query Match 18.4%; Score 33.2; DB 2; Length 186537;
 Best Local Similarity 56.4%; Pred. NO. 19;
 Matches 62; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY 7 GCGATAAGCTATCTGATGTCAGTTGTGACGAACCTCTTTTACACCGAATCGTT 66
 Db 181188 GCGTGTGAGCGATTAATAACCTTTTTCATCTTCACCTTTTCTTTATCTGATTTTC 181247

QY 67 CTGAGTCATCTTTTGGCGGCAATCTATGGAACCTTACCTCATTTTCT 116
 Db 181248 ATTGTATATTTTGTGAGGTCTTATCTAAACCTTTCCCATCTCT 181297

RESULT 14
 AC019181/c
 LOCUS 190998 bp DNA linear PRI 08-NOV-2000
 DEFINITION Homo sapiens BAC clone RP11-272E3 from 2, complete sequence.
 AC019181
 ACCESSION AC019181.4 GI:8748861
 VERSION AC019181.4
 KEYWORDS HTG.
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 190998)
 TATeno.M., Catanese.J.J. and de Jong.P.J. (1998) An improved
 approach for construction of bacterial artificial chromosome
 libraries. Genomics 51:1-8. The clone may be obtained either from
 Research Genetics, Inc. (http://www.resgen.com) or Pieter de Jong
 and coworkers at the Roswell Park Cancer Institute
 (http://bacpac.med.buffalo.edu)
 VECTOR: pBACe3.6
 NEIGHBORING SEQUENCE INFORMATION:
 The clone sequenced to the left is RP11-449E13. Actual start of
 this clone is at base position 1 of RP11-272E3; actual end is at
 base position 190998 of RP11-272E3.

FEATURES
 Location/Qualifiers
 1..190998
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="2"
 /map="2"
 /clone="RP11-272E3"
 /clone_lib="RPCI-11"
 359..475
 /rpt_family="ERV1"
 repeat_region
 859..1355
 /rpt_family="ERV1"
 repeat_region
 3639..3723
 /rpt_family="ERV1"
 repeat_region
 3756..4052
 /rpt_family="Alu"
 repeat_region
 4054..4330
 /rpt_family="Alu"
 repeat_region
 4378..4693
 /rpt_family="ERV1"
 repeat_region
 6155..6190
 /rpt_family="MaLR"
 repeat_region
 6168..6275
 /rpt_family="5S"
 repeat_region
 6257..6616
 /rpt_family="MaLR"
 repeat_region
 7190..7486
 /rpt_family="Alu"
 repeat_region
 7948..8107
 /rpt_family="L1"
 repeat_region
 8288..8583
 /rpt_family="Alu"
 repeat_region
 8585..8923
 /rpt_family="MER2_type"
 repeat_region
 8965..10530
 /rpt_family="MER2_type"
 repeat_region
 10574..10628
 /rpt_family="L1"
 repeat_region
 10676..10962
 /rpt_family="Alu"
 repeat_region
 10966..11487
 /rpt_family="MER2_type"
 repeat_region
 11787..12261
 /rpt_family="L1"
 repeat_region
 12265..12334
 /rpt_family="Alu"

Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site: http://genome.wustl.edu/gsc
 Contact: sapiens@wustl.wustl.edu
 Summary Statistics
 Center project name: H_NH0272E03

NOTICE: This sequence may not represent the entire insert of this
 clone. It may be shorter because we only sequence overlapping
 clone sections once, or longer because we provide a small overlap
 between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:

all regions were double stranded, sequenced with an alternate
 chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such
 as compressions and repeats; all regions were covered by sequence
 from more than one subclone; and the assembly was confirmed by
 restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. John D.
 McPherson, Department of Genetics, Washington University, St. Louis
 MO. For additional information about the map position of this
 sequence, see http://genome.wustl.edu/gsc

SOURCE INFORMATION:

The RPCI-11 human BAC library was made from the blood of one male
 donor, as described by Osoegawa,K., Woon,P.Y., Zhao,B., Frengen,E.,
 Tatenio,M., Catanese,J.J. and de Jong,P.J. (1998) An improved
 approach for construction of bacterial artificial chromosome
 libraries. Genomics 51:1-8. The clone may be obtained either from
 Research Genetics, Inc. (http://www.resgen.com) or Pieter de Jong
 and coworkers at the Roswell Park Cancer Institute
 (http://bacpac.med.buffalo.edu)

VECTOR:

pBACe3.6

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is RP11-449E13. Actual start of
 this clone is at base position 1 of RP11-272E3; actual end is at
 base position 190998 of RP11-272E3.

FEATURES

Location/Qualifiers

1..190998

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="2"

/map="2"

/clone="RP11-272E3"

/clone_lib="RPCI-11"

359..475

/rpt_family="ERV1"

repeat_region

859..1355

/rpt_family="ERV1"

repeat_region

3639..3723

/rpt_family="ERV1"

repeat_region

3756..4052

/rpt_family="Alu"

repeat_region

4054..4330

/rpt_family="Alu"

repeat_region

4378..4693

/rpt_family="ERV1"

repeat_region

6155..6190

/rpt_family="MaLR"

repeat_region

6168..6275

/rpt_family="5S"

repeat_region

6257..6616

/rpt_family="MaLR"

repeat_region

7190..7486

/rpt_family="Alu"

repeat_region

7948..8107

/rpt_family="L1"

repeat_region

8288..8583

/rpt_family="Alu"

repeat_region

8585..8923

/rpt_family="MER2_type"

repeat_region

8965..10530

/rpt_family="MER2_type"

repeat_region

10574..10628

/rpt_family="L1"

repeat_region

10676..10962

/rpt_family="Alu"

repeat_region

10966..11487

/rpt_family="MER2_type"

repeat_region

11787..12261

/rpt_family="L1"

repeat_region

12265..12334

/rpt_family="Alu"

repeat_region


```
gene
CDS
/db_xref="taxon:71421"
165..1403
/gene="H11348"
165..1403
/gene="H11348"
/notes="similar to SP:P29745 GB:U00096 PID:1651556
PID:1787372 percent identity: 54.27; identified by
sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="peptidase T (pepT)"
/protein_id="AAC22995.1"
/db_xref="GI:1574810"
/translation="MISQDKTELLERFLHYVSPHTQSKPNKAKHSPSSVGOMKLAQOL
QKELTQLGLENVEVSKYAVVTFALPANDPNLTGTLVAHLDTSPOCGKNNRVPVIE
EYRGDIALGIEERFISPVYSFMOKLVGQTLIVDGTLLGADNKGACIETMTALSI
LOKENIPCNIRVATPDEEIGLGHYFPEMEKFCSDWAYTDGGVGVGELEVENFNAAT
AKVRFGRNHTGYSARGKMLNALTACCEQVQVFPVDEPEKTDGKVGFGYHLEDFSGDI
EVEVEITYLRDDEQNFQAKRAFIKNQVEKFNKGLKKPIELEIQDSYQNNYDVVN
VQSIELADRAMKAVGIRPNHKPIRGGTDGAPLASKGLACPNIFTGGYNFHSKHELVS
LQGMENFVQVITEMLKCKDL"
complement(1967..2449)
/gene="H11349"
complement(1967..2449)
/gene="H11349"
/notes="similar to PID:1652428 percent identity: 31.37;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAC22996.1"
/db_xref="GI:1574811"
/translation="MSKTSIGLQKVSQSAELADKLNELLATYQVFTNVRGVHWNKGV
NFFALHAKFEETYNLVARVDEVAERILTLGVTNNAYSOYLKRIKEDLAVSAQE
CLSGTLQGLKTLDDQREILAPANNANDEGTASQMSDYIKEQEKLVWFPQACQTCN
2723..3601
/gene="H11350"
2723..3601
/gene="H11350"
/notes="similar to GB:X63144 GB:U15410 GB:X16419 PID:41090
PID:560012 percent identity: 51.18; identified by sequence
similarity; putative"
/codon_start=1
/transl_table=11
/product="cytidine deaminase (cdd)"
/protein_id="AAC22997.1"
/db_xref="GI:1574812"
/translation="MQDLIKRTLPODDALNQAIYVNDLRSONWAGELNYSQVOQLCHNF
EUTLLKLAHLLPLAASYSHTAISHFNVGAIGAIGQDDFYFGANQEFANSAIQOTIHA
EQSAISHAWLRNERRISDMVNYTPCGHCRQPMNELHGAEKISIHLPSONNPLHSYL
PDAGFGPKDLDTAAHLLAEENHDLVADHODLVNQAIILAAQSHCPYSNSPHGIALLFK
NSDVVTGRVAENAFNPSPALQATLNFAYLNDKKLSLSDIERIVMAEKALKLSHKTMAE
TLLSLTSLVELEYSL"
complement(3644..4609)
/gene="H11351"
complement(3644..4609)
/gene="H11351"
/notes="similar to GB:U00096 PID:1788178 percent identity:
65.22; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAC22998.1"
/db_xref="GI:1574813"
/translation="MIDRPFYQIATTNLSDWLETLPCOLKEWETQTHGYAKWSKI
VDFNLNHADEIDLKSAKSDRTSPLSEGEKORIITHLKQLMPKRGYTHLFGIIVDC
EWRSDFKWDRVLPPLSQGRITLIDVGGSGYHMMRWGEGAKMYVGDPLFELFCQF
EAVRLKLNDRANILPIGIEQMPPLAADFVTFMSGLYHHRKSLDHLSQLKNQLVKG
GGLVLETVDDGDIINTVLVADPYAKMKNVYFIPSVATLINNLEKLVGFTNVRCDVAT
TTLEBQRTDWLENSLIDFLDPNDHSKTIIEGYQAPKRAVILANK"
complement(4606..6120)
/gene="H11352"

CDS
complement(4606..6120)
/gene="H11352"
/notes="similar to GB:D32069 PID:509694 PID:913690 percent
identity: 64.98; identified by sequence similarity;
putative"
/codon_start=1
/transl_table=11
/product="sodium/proline symporter (proline permease)
(putP)"
/protein_id="AAC22999.1"
/db_xref="GI:1574814"
/translation="MFGDPDSLITFTIYIFGMLLGLVLAAYVTNNLSVDYLGGRLGGS
FVTNAGASGMSWGLMGLPGAVYLSGLVEGIAIGLTIGAYFNMLLVAGRLRYTE
LNNALTYLPEYFHNFGSSHKLKLVLSATILVFTIYCASCVAGAKLQNFISVEY
STALWALGAATYATYFIFGFLAVSWTDIATIMFALITLTPVLLSFADPAQFSAY
LEOAAVYNKQFTDLFTSTPLGLLSLAAMGLGYGQPHILARFNAADSVKSLIKARR
ISMGMVLCLAGAIGLGAIPYFFANPAIAGTVNREPEQVFIELAKLLFNWPIAGIL
LSAIIAAVAGTSLSAQLLISSSITDEFYKGFIRPNASEKELVNLGRIMVLVIAALAIW
TAODNSKVLKLVPEAWAGFAGPVVLFSLFWRMTSSGAMGLMGLGAVTVFAMKE
VVPADTFKVVYEMIPGPAFASLAIIIVISLSNKKPEQDILNTFDKAEAKYKAEK"
6236..7711
/gene="H11353"
6236..7711
/gene="H11353"
/notes="similar to SP:P25537 PID:48826 PID:606187 GB:U00096
PID:1789645 percent identity: 70.96; identified by
sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="cytoplasmic axial filament protein (cafa)"
/protein_id="AAC23000.1"
/db_xref="GI:1574815"
/translation="MDAVELLMNVTPNETRIALVETGMLREVHIERQAKRGIVGNIYK
GRVTLPMQOSAFVDIGLEAAFLHAAIDIVSHTECVDENEOKQFKVKSISLVRREGO
DIVGVVKEPTGKCARLTDTITLSRHLVPENSHVGYQSRISEERARLKAIVE
PFCDEGLIIRKATEGASEELRODEFLKARLWRKRVLERKSKYPTKSIYGEIPALPO
RILRDFIGNLLEKIRIDSKLFGCEKFTDEPMELSKLVLSQNGPIFDYGVVNA
IQTALDRVNLKSGYLIEQTEAMTTIDINTGAFVGHRLNEETIFNTIEATKAIAH
EQLRLNLGIIIDIDMOTDEHRNRVLSLCLDLSKDRMTNNGFTQLGLVEMTRK
RTRESLHVLCDECTPCGGRGVKVTVCYIEMREIRVYHLSFSESEQFVVYASPAVS
EYLINESHGLLPEVEMFGRKRVKTEQFYNQEQFDVVM"
8131..9804
/gene="H11354"
8131..9804
/gene="H11354"
/notes="similar to GB:V01575 SP:P00962 GB:J01617 GB:M0187
PID:146168 percent identity: 76.28; identified by sequence
similarity; putative"
/codon_start=1
/transl_table=11
/product="glutaminyl-tRNA synthetase (glns)"
/protein_id="AAC23001.1"
/db_xref="GI:1574816"
/translation="MMSHTETSLGAENTRTHNFITQIIDEDLASCKHKSVMHFRFPPEP
NGYLHGHAKSICLNFLGAKETQGCNLRFDTPNVKEDVEYVDSIKADVWLGFKWE
GEPYASDYDALYGYAVLKKGLAYVDELSPEMREYRGLTFEGKNSPYRDRITIE
ENLALFERKNGEFAEGKASRAKIDMASPEMREPVYIRKFSHHQTKDWKCIYP
MYDFTHCISDAIETHSTCTLEFQDNRLYDWLENISIERLPHQVFSRLNLEGT
LTSKRLLKLVNDEIVDGNWDPMPITISLRRRGYTPASLREFCRIGVTKODNVVEY
SALEACIRLEDNLNAPRAMVIDPVYVNIENFSEAVLTAPNHPPELGERQLPFTK
ELIYDRADFREANKQYKRLVGLGKVRANAYIKAEKVEKDANGSEITITICTYDET
LGKNPADGRKGVKGIHWVSANNHHPAEFLYDLRFTVPNPGAEDDIESVLPNSLVIK
QGFVEQSLANAEAKGYQFEREGYFCADSKOSRPEHLVNLTVSLKEGF"
9870..10334
/gene="H11355"
9870..10334
/gene="H11355"

Query Match
Best Local Similarity 18.1%; Score 32.6; DB 1; Length 10393;
Matches 65; Conservative 0; Mismatches 54; Indels 0; Gaps 0;
99 GAATCGTCTGAGTCATCTTTTGGCGGCGGAATCTATGGGAACCCCTACCTCATTTCTCC 118
```

Db 8136 GAGCCAACAGAGACATCTTTGGGTGCAGAAATACCCGCACTCATAATTTTATACCCA 8195
Oy 119 AGACATCGCTCCGAAATTACCTCCGCGAACCTACTATATCTTCTCCTCGGTTTCCTC 177
Db 8196 AATTATGATGAAGATTAGCTTCGGGTAACATAAAAGCGTTTCATACCCGTTTTCGCG 8254

Search completed: June 6, 2003, 15:13:59.
Job time : 529.302 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nuclei: search, using sw model

Run on: June 6, 2003, 12:27:48 ; Search time 467.887 Seconds
(without alignments)
6230.541 Million cell updates/sec

Title: US-09-775-879-20_COPY_143_322
Perfect score: 180
Sequence: 1 a1agcgcggaatgctta.....tcattctcgggttctctgg 180

Scoring table: IDENTITY_NUC
Gap 10.0, Gapext 1.0

Searched: 161:4066 seqs, 8097743376 residues
Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST *

- 1: em_estba:**
- 2: em_esthum:**
- 3: em_estin:**
- 4: em_estmu:**
- 5: em_estov:**
- 6: em_estpl:**
- 7: em_estro:**
- 8: em_htc:**
- 9: gb_est1:**
- 10: gb_est2:**
- 11: gb_htc:**
- 12: gb_est3:**
- 13: gb_est4:**
- 14: gb_est5:**
- 15: em_estfun:**
- 16: em_estom:**
- 17: gb_gss:**
- 18: em_gss_hum:**
- 19: em_gss_inv:**
- 20: em_gss_pln:**
- 21: em_gss_vrt:**
- 22: em_gss_fun:**
- 23: em_gss_mam:**
- 24: em_gss_mus:**
- 25: em_gss_othr:**
- 26: em_gss_pro:**
- 27: em_gss_rod:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	88.4	49.0	551	10	BE471892	BE471892 EST416745
2	88.4	49.0	597	12	BG098248	BG098248 EST462767
3	88.4	49.0	572	12	BG600842	BG600842 EST505737
4	87.8	48.0	572	10	BE471514	BE471514 EST416367
5	87.8	48.0	608	14	BQ118369	BQ118369 EST603945
6	86.8	48.0	513	10	AW429091	AW429091 EST306463

7	85.2	47.3	753	13	BI930355	BI930355
8	83.8	46.6	621	12	BG123510	BG123510
9	83.6	46.4	567	10	AW429266	AW429266
c 10	76.4	42.4	486	14	BQ028052	BQ028052
11	76	42.2	535	10	AW596303	AW596303
12	76	42.2	583	12	BF425767	BF425767
13	75	41.7	722	12	BG598293	BG598293
14	71.2	39.6	444	13	BI470419	BI470419
15	70.8	39.3	535	13	BM522429	BM522429
16	69.6	38.7	454	10	AW624414	AW624414
17	69.6	38.7	458	9	AU163003	AU163003
18	68.8	38.2	592	10	AW219366	AW219366
19	67.6	37.6	359	9	AU068067	AU068067
20	67.6	37.6	458	9	AU070031	AU070031
21	58	32.2	405	12	BG487492	BG487492
22	58	32.2	541	12	BG322635	BG322635
23	58	32.2	562	10	BE125051	BE125051
24	58	32.2	570	13	BG933493	BG933493
25	58	32.2	587	12	BF586850	BF586850
26	58	32.2	667	10	BE357415	BE357415
27	55.4	30.8	530	13	BJ200021	BJ200021
28	54.8	30.4	604	13	BJ195304	BJ195304
29	54.8	30.4	607	13	BJ202813	BJ202813
30	54.8	30.4	645	13	BJ157215	BJ157215
31	54.6	30.3	804	13	BI308419	BI308419
c 32	53.2	29.6	252	17	BH215130	BH215130
33	53.2	29.6	605	13	BI417093	BI417093
34	53.2	29.6	605	13	BI674125	BI674125
35	52.8	29.3	716	10	AW75288	AW75288
36	52	28.9	825	10	BE643026	BE643026
37	50.4	28.0	586	14	BM903391	BM903391
38	48	26.7	517	13	BJ188385	BJ188385
39	44.6	24.8	483	13	BJ199974	BJ199974
40	44.6	24.8	492	13	BJ197093	BJ197093
41	44.6	24.8	517	14	BQ703739	BQ703739
42	42	23.3	306	12	BF483844	BF483844
43	42	23.3	409	9	AL819228	AL819228
44	42	23.3	658	14	BQ246840	BQ246840
45	40.6	22.6	325	12	BG265811	BG265811

ALIGNMENTS

RESULT 1: BE471892 551 bp mRNA linear EST 28-JUL-2000
LOCUS BE471892
DEFINITION EST416745 potato stolon, Cornell University Solanum tuberosum cDNA clone CSTA30N16, mRNA sequence.
ACCESSION BE471892
VERSION BE471892
KEYWORDS EST.
SOURCE Potato.
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
REFERENCE 1 (bases 1 to 551)
AUTHORS van der Hoeven, R., Bezzerides, J., Bachem, C., Horvath, B., Visser, R., Holt, I.E., Liang, F., Hansen, T.S., Utterback, T., Bowman, C.L., Doan, B., Bougri, O., Buell, C.R., Ronning, C.M., Tanksley, S.D. and Baker, B.
TITLE Generation of ESTs from potato swelling stolons
JOURNAL Unpublished (1999)
COMMENT Contact: Research Genetics, Libraries Division
Tel: 1-800-711-6195
Email: cdna@resgen.com
5 prime sequence.
FEATURES
source Location/Qualifiers
1..551
/organism="Solanum tuberosum"
/cultivar="Bintje"
/db_xref="taxon:4113"

BASE COUNT	155 a	138 c	125 g	179 t	frozen in liquid nitrogen immediately upon removal from tubers."
ORIGIN					
Query Match		49.1%	Score 88.4;	DB 12;	Length 597;
Best Local Similarity		68.5%;	Pred. No. 1.6e-17;		
Matches 122;	Conservative	0;	Mismatches 56;	Indels 0;	Gaps 0;
Qy	3	GGCGGGCGGATAATGCTTATCTGATGCGAGTTGTGTGACGAAACCTCTTTTTACAAACCGAAT	62		
Db	70	GTCCGAGATCGAGGATTACTTGAAGCTGTTCGTGGAGGAGACATCGTTTACAAACCGTAT	129		
Qy	63	CGTTCGTGAGTCATCTTTTGGCGGGGAATCATATGGGAACCCCTTACCTCATTTTCTCCAGAC	122		
Db	130	GGTTTTGGGFPACATTTCTTGGCCGGAATCATCGTGGGGACCCTTCTCTCATATGCTTCAAGG	189		
Qy	123	ATGGCTCCGAAATTAACCTCCGCCGGAACCCCTACATATCTTCACTCCGGTTTCTCTCG	180		
Db	190	ATGGCTCCGFPAACTACATATGGCGGTGTTTACTTTACTTCTCATCTCCGGTTTCTCTCG	247		

RESULT	3
BG600842	
LOCUS	
DEFINITION	BG600842 672 bp mRNA linear EST 12-APR-2001 EST505737 cSTS Solanum tuberosum cDNA clone cSTS300I3 5' sequence, mRNA sequence.
ACCESSION	BG600842
VERSION	BG600842.1 GI:13617978
KEYWORDS	EST.
SOURCE	potato.

LOCUS	BG600842	672 bp	mRNA	linear	EST 12-APR-2001
DEFINITION	EST505737 csts	Solanum tuberosum	cdna clone	cSTS300l3	5' sequence,
					mRNA sequence.

BG000642
 BG600842.1 GI:13617978
 EST.
 KEYWORDS
 SOURCE
 potato.
 Solanum tuberosum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
 1 (bases 1 to 672)
 van der Hoeven,R., Bezzerides,J., Sun,H., Cho,J., Chiemiango,A.,
 Bougri,O., Buell,C.R., Ronning,C., Tanksley,S. and Baker,B.
 Generations of ESTs from sprouting potato eyes
 Unpublished (2000)
 Contact: Cathy Ronning

```

ORGANISM
Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
1 (bases 1 to 672)
van der Hoeven,R., Bezzerides,J., Sun,H., Cho,J., Chiemingo,A.,
Bougri,O., Buell,C.R., Ronning,C., Tanksley,S. and Baker,B.
Generations of ESTs from sprouting potato eyes
Unpublished (2000)
Contact: Cathy Ronning
The Institute for Genomic Research
For clone info: please contact Research Genetics, Libraries
Division tel 1-800-711-6195, email cdna@resgen.com
Seq primer: M13F-R.
Location/Qualifiers
1..672
/organism="Solanum tuberosum"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="CSTS30013"
/clone_lib="cSTS"
/tissue_type="sprouting eyes from tubers"
/dev_stage="12-14 weeks post harvest"
/lab_host="SOLR"
/note="vector"
XhoI: Various sizes of sprouting eyes (2mm to 15mm) were
taken from tubers. The tubers were incubated at 26C in the
dark for 2-3 weeks prior to sprouting. The eyes were
frozen in liquid nitrogen immediately upon removal from
tubers."
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
source

```

```
|||||
171 GGTGTTTGGTACATCTTGGCCGAATCATGTTGGGACCACTTCCTCATATGCTTCAAGG 230
LOCUS
DEFINITION
123 ATGGCTCGGAATTAACCTCGCGGAACCTTACTATCTCTCGGTTTCTCTGG 180
Db
231 ATGGCTCGTAACACTACATGGCGGTGTTTACTTTACTTCACTCCGGTTTCTCTGG 288

RESULT 4
BE471514
LOCUS
DEFINITION
Estr416367 potato stolon, Cornell University Solanum tuberosum cDNA
clone csta29120, mRNA sequence.
BE471514
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
REFERENCE
1 (bases 1 to 572)
AUTHORS
van der Hoeven,R., Bezzerides,J., Bachem,C., Horvath,B., Visser,R.,
Holt,I.E., Liang,F., Hansen,T.S., Utterback,T., Bowman,C.L., Doan
,B., Bouffri,O., Buell,C.R., Ronning,C.M., Tanksley,S.D. and Baker
,B.
TITLE
Generation of ESTs from potato swelling stolons
JOURNAL
Unpublished (1999)
CONTACT: Research Genetics, Libraries Division
Tel: 1-800-711-6195
Email: cna@resgen.com
5 prime sequence.
FEATURES
Location/Qualifiers
1..572
/organism="Solanum tuberosum"
/cultivar="Bintje"
/db_xref="taxon:4113"
/clone="csta29120"
/tissue_type="potato stolon, Cornell University"
/stolons="axillary buds of stem explants, swelling
stolons"
/dev_stage="1 to 3 days"
/lab_host="SOLR"
/Note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; RNA was supplied by Christian Bachem & Beatrix
Horvath(Laboratory of Plant Breeding, Dept. of Plant
Sciences, Wageningen University, The Netherlands). Total
RNA was isolated from developing axillary buds of potato
nodal stem cuttings cultured on medium for the
introduction of tuber formation as described in Bachem et
al. (Plant Journal 1996). Tissue samples were taken of
stages corresponding to growing stolons and the early
stages of tuber formation."
BASE COUNT 135 a 137 c 122 g 178 t
ORIGIN
Query Match 48.8%; Score 87.8; DB 10; Length 572;
Best Local Similarity 68.2%; Pred. No. 2.5e-17;
Matches 122; Conservative 0; Mismatches 57; Indels 0; Gaps 0;
QY 2 TGGCGCGGATAATCGTTATCTGATGCGAGTTGTTGGAGAAACCTCTTTTACACCGAA 61
|||||
Db 113 TCTCGCAGATGAGGACTTACTTGAAGCTGTCGTGGAGGAGACTTCGTTTACACCGTC 172
QY 62 TCGTTC TGAGTCATCTTTTGGCGGGAATCTATGGAACCTTACCTCATTTTCTCCAGA 121
|||||
Db 173 TGGTTT TGGGTACATCTTCTGCGCGGAATCATGTGGGGACCACTTCTCATATGCTTCAAG 232
QY 122 CATGCT TCGGAATTAACCTCGCGGAACCTACTATCTATCTATCTCCGGTTTCTCTGG 180
|||||
Db 233 GATGGC TCGGTAACTACATGCGGGTGTTTACTTACTTCACTCATCTCCGGTTTCTCTGG 291
```

```
RESULT 5
BQ118369
LOCUS
DEFINITION
EST603945 mixed potato tissues Solanum tuberosum cDNA clone STMEA40
5' end, mRNA sequence.
BQ118369
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
REFERENCE
1 (bases 1 to 608)
AUTHORS
Buell,C.R., Hart,A., Baker,B., Tanksley,S., Fry,W., Smart,C.,
Restrepo,S., Griffiths,H., van der Hoeven,R., Tsai,J. and
Karanycheva,S.A.
TITLE
Generation of a set of potato cDNA clones for microarray analyses
JOURNAL
Unpublished (2002)
COMMENT
On Apr 17, 2002 this sequence version replaced gi:20170331.
CONTACT: Robin Buell
The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato@tigr.org
This clone is available through the Research Genetics, contact the
Research Genetics for further information 1-800-711-6195 or
cna@resgen.com
Seq primer: T3
FEATURES
Location/Qualifiers
1..608
/organism="Solanum tuberosum"
/cultivar="Kennebec or Binjfe"
/db_xref="taxon:4113"
/clone="STMEA40"
/clone_lib="mixed potato tissues"
/tissue_type="mixed tissues"
/lab_host="SOLR"
/Note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; supplier: Combination of untreated and Phytophthora
infestans-treated libraries of stolons, leaves, leaflets,
axillary buds of stem explants, petioles, germinating eyes
tubers, or roots."
BASE COUNT 155 a 144 c 125 g 184 t
ORIGIN
Query Match 48.8%; Score 87.8; DB 14; Length 608;
Best Local Similarity 68.2%; Pred. No. 2.5e-17;
Matches 122; Conservative 0; Mismatches 57; Indels 0; Gaps 0;
QY 2 TGGCGCGGATAATGCTTATCTGATGCGAGTTGTTGACGAAACCTCTTTTACACCGAA 61
|||||
Db 86 TGTGCGAGATGGAGGATTACTTGAAGCTGTCGTGGAGGAGACTTCGTTTACACCGTC 145
QY 62 TCGTTC TGAGTCATCTTTTGGCGGGAATCTATGGAACCTTACCTCATTTTCTCCAGA 121
|||||
Db 146 TGGTTT TGGGTACATCTTTCGCGGAATCATGTGGGGACCACTTCTCATATGCTTCAAG 205
QY 122 CATGCT TCGGAATTAACCTCGCGGAACCTACTATCTATCTCCGGTTTCTCTGG 180
|||||
Db 206 GATGGC TCGGTAACTACATGCGGGTGTTTACTTACTTCACTCATCTCCGGTTTCTCTGG 264

RESULT 6
AW429091
LOCUS
DEFINITION
EST306463 tomato flower buds 0-3 mm, Cornell University
Lycopersicon esculentum cDNA clone ctoA2A22 5', mRNA sequence.
AW429091
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
tomato.
Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
```



```

Db      310 TTCGTCACAAAGAACCTCTTTTCACACCAAAATAGTGTCTGGAACTTTACTGCCGGAACAA 251
QY      91 CTAATGGAAACCTTACCTCATTTCTCCAGACATGGCTCCGAAATATACCTCCCGGAACC 150
        || || || || || || || || || || || || || || || || || || || || || ||
Db      250 CTCAGACCCGACTCCCGCACCCCTTCAAACTGGCTCCGTAACATTCGCGGATCC 191
QY      151 CTACTATATCATCTCCGGTTTCTCTCTGG 180
        || || || || || || || || || || || || || || || || || || || || || ||
Db      190 CTGTTTACCTCATCTCTGGCTTCCTTTGG 161

RESULT 11
AW596303 535 bp mRNA linear EST 03-DEC-2001
sj01f11.y1 Gm-c1032 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
Gm-c1032-646 5' similar to TR:Q9ZT29 Q9ZT29 STEROL-C5(6
)-DESATURASE. ;, mRNA sequence.
AW596303 1 GI:7283699
EST.
SOURCE soybean.
ORGANISM Glycine max

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.
1 (bases 1 to 535)
Shoemaker,R., Keim,P., Vodkin,L., Erpellding,J., Coryell,V., Khanna
,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,
Wyllie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers
,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk
,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann
,R., Waterston,R. and Wilson,R.
Public Soybean EST Project
Unpublished (1999)
Contact: Shoemaker R/Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available through: ResGen, Invitrogen Corp. 2130
South Memorial Parkway Huntsville, AL 35801 For further information
call: (800)-533-4363 or contact via email: ccu@resgen.com
Insert Length: 1224 Std Error: 0.00
High quality sequence stop: 380.

FEATURES
        source
        1..535
        /organism="Glycine max"
        /db_xref="taxon:3847"
        /clone="GENOME SYSTEMS CLONE ID: Gm-c1032-646"
        /clone_lib="Gm-c1032"
        /tissue_type="Cotyledons of 8-day-old 'Williams'
        seedlings"
        /lab_host="DH10B"
        /note="Vector: pBluescript II XR; Site:1: EcoRI; Site:2:
        XhoI; This cDNA library was constructed from mRNA isolated
        from cotyledons of 8-day-old 'Williams' seedlings which
        were propagated on paper towels with distilled water for
        3 days (etiolated), then greenhouse grown for 5 days in
        potting soil. The cotyledons were flash-frozen in liquid
        nitrogen. StrataGene's cDNA Synthesis Kit (catalog number
        200401) was used to synthesize the cDNA. First-strand
        synthesis was performed with 5-methyl dCTP, hence the
        cDNA was hemimethylated. A modification of
        StrataGene's first-strand synthesis primer was used. An
        'anchor' nucleotide (V-A, C, or G) was added to the 3' end
        of the primer [GAGAGAGAGAGAGAGAACTGTCAG(T)18V] to
        anchor the primer at the 5' end of the poly(A) tract.
        After second strand synthesis, the cDNA ends were filled
        in with cloned pfu DNA, ligated to EcoRI adapters and
        subsequently phosphorylated. The cDNA was then

```

```

precipitated and redissolved in sterile, RNase-,
DNase-free water. The xhoI site within the first-strand
synthesis primer was then restricted by digestion with
XhoI from Promega (400/ul); all xhoI sites in the cDNA
would be protected by their hemimethylated status. The
cDNA constructs were size-fractionated with a 500 bp
cutoff, using GibcoBRL Life Technologies' cDNA Size
Fractionation column. The column eluent was then
precipitated, redissolved, and ligated into StrataGene's
pBluescript II XR Predigested vector (pBluescript II SK(+))
vector that has been digested with EcoRI and XhoI, and
phosphorylated by StrataGene). 100% of the white and 87.5%
of the blue colonies appear to contain recombinant
plasmids with cDNA inserts, based on size (n=28 and 8
respectively). This library was constructed by Dr. Paul
Keim and Dr. Virginia Coryell."
BASE COUNT 121 a 138 c 110 g 163 t 3 others
ORIGIN

Query Match 42.2%; Score 76; DB 10; Length 535;
Best Local Similarity 65.1%; Pred. No. 1.4e-13;
Matches 112; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

QY      9 GGATATGCTTATCTGATGTCAGTTTGTTCAGAAACCTCTTTTACAAACCGAATCGTTCT 68
        || || || || || || || || || || || || || || || || || || || || || ||
Db      1 GAACAGCGAATATTCGAGGCTGTTCGTCGAGGACACGAGACTTGTACAATCGCATTTGTT 60
QY      69 GAGTCATCTTTTTCGCGGCGAATCTATGGAAACCTTACCTCATTTTCTCCAGACATGGCT 128
        || || || || || || || || || || || || || || || || || || || || || ||
Db      61 GGAACCCCTCTTTCGCGACACTTCGTGTGGGGCCACTCCGCGACTTCTCCAGACATGGCT 120
QY      129 CGGAATTAACCTCGCGGAAACCTACTATATCTATCTCCGGTTTCTCTCGG 180
        || || || || || || || || || || || || || || || || || || || || || ||
Db      121 TCGCAATTACCTCGCGGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTGG 172

RESULT 12
BF425767 583 bp mRNA linear EST 06-DEC-2001
sr42f04.y1 Gm-c1051 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
Gm-c1051-728 5' similar to TR:Q9ZT29 Q9ZT29 STEROL-C5(6
)-DESATURASE. ;, mRNA sequence.
BF425767 1 GI:11413756
EST.
SOURCE soybean.
ORGANISM Glycine max

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.
1 (bases 1 to 583)
Shoemaker,R., Keim,P., Vodkin,L., Erpellding,J., Coryell,V., Khanna
,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,
Wyllie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers
,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk
,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann
,R., Waterston,R. and Wilson,R.
Public Soybean EST Project
Unpublished (1999)
Contact: Shoemaker R/Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available through: ResGen, Invitrogen Corp. 2130
South Memorial Parkway Huntsville, AL 35801 For further information
call: (800)-533-4363 or contact via email: ccu@resgen.com
High quality sequence stop: 393.

FEATURES
        source
        1..583

```


by Dr. J. Specht, University of Nebraska (Shoemaker and Specht, 1995). The cDNA library was constructed from mRNA isolated from leaf tissue at various developmental stages of 3 week old greenhouse grown plants. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(dT) sequence with a XhoI restriction site and a 3' anchor. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into DH10B host cells (GibcoBRL). The library was constructed in cooperation with Dr. Paul Keim's laboratory at Northern Arizona University."

BASE COUNT 97 a 103 c 99 g 145 t

Query Match 39.6%; Score 71.2; DB 13; Length 444;

Best Local Similarity 63.4%; Pred. No. 4.3e-12;

Matches 109; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 9 GGATATGCTTATCTGATGCTGTTGTCGAGAAACCTCTTTTACACCGAATCTTCT 68

DB 49 GRACACGAGATATCGAGCTGTTGCGGAGACACGAGACTTGACACCGCATGTGTT 108

QY 69 GAGTATCTTTTCCGCGGGAATCTATGGAACCCCTTACCTCATCTTCTCCAGACATGGCT 128

DB 109 GGGCGCCCTATTGCCGCACTCTGTGTGGGGCCCACTCCGCGCTTCTCCAGACGTGGCT 168

QY 129 CCGAATTACCTCGCGGACCCCTACTATACCTTCATCTCCGGTTCCTCTGG 180

DB 169 TCGCAATTACCTCGCGGCGTCTCTCTACCTCTCTCTGGGCTTCTGG 220

RESULT 15

BM522429

DEFINITION sak98h05.y1 Gm-c1057 Glycine max cDNA clone SOYBEAN CLONE ID:

Gm-c1057-3826 5' similar to TR:Q9SWU4 Q9SWU4 DELTA7 STEROL C-5

DESATURASE. [2] TR:Q39208 ;, mRNA sequence.

BM522429

BM522429.1 GI:18709477

EST.

soybean.

Glycine max

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;

Glycine.

1 (bases 1 to 535)

Shoemaker, R., Keim, P., Vodkin, L., Erpelding, J., Coryell, V., Khanna

, A., Bolla, B., Maria, M., Hillier, L., Kucaba, T., Martin, J., Beck, C.,

Wyllie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers

, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk

, R., Ritter, E., Kohn, S., Shin, T., Jackson, I., Cardenas, M., McCann

, R., Waterston, R. and Wilson, R.

Public Soybean EST Project

Unpublished (1999)

Contact: Shoemaker R/Public Soybean EST Project

Public Soybean EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

This clone is available through: Resgen, Invitrogen Corp. 2130

South Memorial Parkway Huntsville, AL 35801 For further information

call: (800)-533-4363 or contact: ccu@resgen.com web site:

www.resgen.com

Seq primer: -40RP from Gibco

High quality sequence stop: 451.

FEATURES

source

1. .535

/organism="Glycine max"
/db_xref="taxon:3847"
/clone="SOYBEAN CLONE ID: Gm-c1057-3826"
/clone_lib="Gm-c1057"
/tissue_type="Degenerating cotyledons, 2 week old
seedling"

/lab_host="DH10B"
/note="Vector: pBluescript II SK+; Site_1: EcoRI; Site_2:
XhoI; The cDNA library was constructed from mRNA isolated
from degenerating cotyledons of 2 week old seedlings from
PI468916. Complementary DNA was synthesized from mRNA
using a primer consisting of a poly(dT) sequence with a
XhoI restriction site. EcoRI adapters were ligated to the
blunt-ended cDNA fragments followed by XhoI digestion.
The cDNA fragments were directionally cloned into the
EcoRI-XhoI restriction site of the pBluescript vector. The
ligated cDNA fragments were transformed into DH10B host
cells (GibcoBRL). This library was constructed in the
laboratory of Dr. Randy Shoemaker."

BASE COUNT 121 a 116 c 110 g 188 t

Query Match 39.3%; Score 70.8; DB 13; Length 535;

Best Local Similarity 64.8%; Pred. No. 6.1e-12;

Matches 105; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

QY 19 TATCTGATGCACTTTGTTGACGAAACCTCTTTTACAACGAATCTTCTGAGTCATCTT 78

DB 38 TATTCGAGCTGTTGCGGCAAGACACGAGACTGTACACCGCATTTGTGTGGCGCCTA 97

QY 79 TTGCGGCGCAATCTATGGAACCCCTTACCTCATCTTTTCTCCAGACATGGCTCCGAAATTAC 138

DB 98 TTGCGGCACTGTGTGTGGGGCCCACTCCCGCGCTTCTCCAGACGTGGCTTCGCAATTAC 157

QY 139 CTCGCGGAACCTACTATATCTCTCCGGTTTCTCTCTGG 180

DB 158 CTCGCGGCGGTTCCTCTACCTCTCTCTGGCTCTTGTGG 199

Search completed: June 6, 2003, 16:23:08

Job time : 477.887 secs

Db 121 ACATGGCTCCGAAATTACCTGCCCGGAACCTACTATCTCATCTCCGGTTCTCTGG 180


```
NAME/KEY: misc_feature
LOCATION: (45593)..(45593)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (45732)..(45732)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (47036)..(47036)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51334)..(51334)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51602)..(51602)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51786)..(51786)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51805)..(51805)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (55369)..(55369)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (65309)..(65309)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (65313)..(65313)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (80024)..(80024)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (100091)..(100091)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (102696)..(102696)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (105121)..(105121)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (107248)..(107248)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (117136)..(117136)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (119750)..(119750)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (119924)..(119924)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
```

```
LOCATION: (120038)..(120038)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (121344)..(121344)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (122167)..(122167)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (122336)..(122336)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (131340)..(131340)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (131360)..(131360)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (139910)..(139910)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (140398)..(140398)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (142750)..(142750)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145058)..(145058)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145171)..(145171)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145942)..(145942)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (147197)..(147197)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (150841)..(150841)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (152500)..(152500)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (152530)..(152530)
```

```
Query Match      18.1%; Score 32.6; DB 9; Length 1830121;
Best Local Similarity 54.6%; Pred. No. 6,6;
Matches 65; Conservative 0; Mismatches 54; Indels 0; Gaps 0;
```

```
Qy 59 GAATCGTTCTGAGTCATCTTTTGGCGGGAATCTATGGGAACCTTACCTCATTTTCCTCC 118
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1431803 GAGCCATACAGAGACATCTTTTGGTGCAGAAAATACCCGCACTCATATAATTTTATTACCCA 1431862
Qy 119 AGACATGCTCCGGAATTAACCTCGCCGGAACCTTACTACTTCTATCTCCGGTTTCCTC 177
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1431863 AATTATTGATGAAGATTTAGCTTCGGGTAAACATAAAAGCGTTTCATACCCGTTTTCCTCCG 1431921
```

RESULT 7
US-09-764-877-3680
: Sequence 3680, Application US/09764877

```

RESULT 8
US/09-880-107-3709
; Sequence 3709, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherif, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression P
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,3
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,0
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3709
; LENGTH: 3810
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession
US/09-880-107-3709

```

	Query Match	16.4%	Score 29.6;	DB 10;	Length 3810;
	Best Local Similarity	51.5%;	Pred. No. 4.7;	Mismatches 64;	Indels 0; Gaps 0;
	Matches	68;	Conservative 0;		
Qy	17	CTTATCTGATGCAGTTTGTTGACGAACCTCTTTTTACAAACCAGTCGTCTCTCAGTCATC	76		
Dd	1203	CTTTTCTGATGGGAATTCTCAAGCGCTCAGTCAGTATATAAGTCACAAGAGCTCAGTGACC	1362		
Qy	77	TTTWTGCCGGGGAATCTATGGGAACCCCTACTCATTTTCTCCAGACATGCGTCCGGAATT	136		
Dd	1263	CAAGGCCAAGATGCTAGTGCTCTCTTANGGCAATGTTCTTTCCTACTAGTACTAGGCT	1322		
Qy	137	ACCTCGCCGGAA	148		
Dd	1323	GCCTCACAGGAA	1334		

```

RESOLUT 12
US-09-923-876-5279
; Sequence 5279, Application US/09923876
; Patent No. US20020013958A1
; GENERAL INFORMATION:
; APPLICANT: Lalqudi, Raghunath V.
; APPLICANT: Kamigaki, Laura Y. (Ito)
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN SEEDLING
; FILE REFERENCE: PL-0012-1 CON
; CURRENT APPLICATION NUMBER: US/09/923,876
; CURRENT FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: 09/298,329
; PRIOR FILING DATE: 1999-04-21
; PRIOR APPLICATION NUMBER: 60/085,331
; PRIOR FILING DATE: 1998-05-05
; NUMBER OF SEQ ID NOS: 6332
; SOFTWARE: PERL Program
; SEQ ID NO 5279
; LENGTH: 277
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Invent ID No. US20020013958A1 700456730H1

```

1


```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on:      June: 6, 2003, 13:35:44 ; Search time 31.9245 Seconds
              (without alignments)
              1729.135 Million cell updates/sec

Title:      US-(9-775-879-20_COPY_143_322
Perfect score: 180
Sequence:    1 atgcgcggcgataatgctta.....tcactcgcggtttctctcgg 180

Scoring table:  IDENTITY_NUC
                  gapcp 10.0 , Gapext 1.0

```

```

RESULT 1
US-09-443-041A-13
; Sequence 13, Application US/09443041A
; Patent No. 6465717
; GENERAL INFORMATION:
; APPLICANT: Famodu, Omolayo O.
; APPLICANT: Orusco, Buddy
; APPLICANT: Rafalski, Antoni
; APPLICANT: Shen, Jennie
; TITLE OF INVENTION: Sterol Metabolism
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/09/443
; CURRENT FILING DATE: 1999-11-18
; PRIOR APPLICATION NUMBER: 60/109,283
; PRIOR FILING DATE: 1998-11-20
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 13
; LENGTH: 459
; TYPE: DNA
; ORGANISM: Glycine max
US-09-443-041A-13

```

RESULT 2
US-09-443-041A-27
: Sequence 27, Application US/09443041A
: Patent No. 6465717
: GENERAL INFORMATION:
: APPLICANT: Famodu, Omolayo O.
: APPLICANT: Orszco, Buddy
: APPLICANT: Rafalski, Antoni
: APPLICANT: Shen, Jennie
: TITLE OF INVENTION: Sterol Metabolism

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

***** DE SEC LENGTH: 2000000000

FOST PROCESSING: MINIMUM MATCH 0%
MAXIMUM MATCH 100%

Listing first 45 summaries

```
Database :
Issued_patents_NA: *
1: /can2 6 /patdata=1 /inc/53 COMB seq +
```

```
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_5/ptodata/1/ina/5A_COMB.seq:*
4: /cgn2_5/ptodata/1/ina/5B_COMB.seq:*
```

```
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
```

```
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*
```

Pred. No.	Is tie number of results predicted by chance to
1	1
2	2
3	3
4	4
5	5
6	6
7	7
8	8
9	9
10	10
11	11
12	12
13	13
14	14
15	15
16	16
17	17
18	18
19	19
20	20
21	21
22	22
23	23
24	24
25	25
26	26
27	27
28	28
29	29
30	30
31	31
32	32
33	33
34	34
35	35
36	36
37	37
38	38
39	39
40	40
41	41
42	42
43	43
44	44
45	45
46	46
47	47
48	48
49	49
50	50
51	51
52	52
53	53
54	54
55	55
56	56
57	57
58	58
59	59
60	60
61	61
62	62
63	63
64	64
65	65
66	66
67	67
68	68
69	69
70	70
71	71
72	72
73	73
74	74
75	75
76	76
77	77
78	78
79	79
80	80
81	81
82	82
83	83
84	84
85	85
86	86
87	87
88	88
89	89
90	90
91	91
92	92
93	93
94	94
95	95
96	96
97	97
98	98
99	99
100	100

and is derived by analysis of the total score distribution

SUMMARIES

Result	Quer?
--------	-------

[illegible]

2	71.2	1447	4	US-09-443-041A-27
2	39.15	1447	4	US-09-443-041A-27

3	09-09-443-041A-23
4	US-09-443-041A-9

6	37.4	20.11	495	4	US-09-443-041A-15
7	38.3	21.11	360	4	US-09-443-041A-II

7	37.7	ZO:II	IZIO	4	US-09-443-04IA-29	See
8	28.8	16.0	2022	1	US-08-803-973-6	See

	Year	US-08-803-972-6	US-08-803-973-17
3	2022	16.0	16.0
10	2024	28.8	28.8

E	11	28.8	16.0	2124	1	US-08-803-972-11	see
C	12	28.8	16.0	4371	1	US-08-803-972-1	see

Case	Age	Sex	Ref	Year	Country	Case	Age	Sex	Ref	Year	Country		
13	28.8	M	16.0	4371	1	US-08-803-972-1	14	28.2	M	15.5	97350	2	US-08-781-801-70

[illegible]

C	17	28	15.€	35100	2	US-08-770-379-19	Sec
c	17	28	15.€	35100	2	US-08-770-379-19	Sec

C	19	28	15.6	35100	4	US-09-230-371A-19	Sec
C	19	28	15.6	35100	4	US-09-230-371A-19	Sec

21	27	15.C	726	4	US-08-896-164-72	Sec
----	----	------	-----	---	------------------	-----

C 23	26.8	14.5	1387	4	US-09-461-474-1	Sec
------	------	------	------	---	-----------------	-----

25	25.6	14.2	4800	2	US-08-612-734B-3	Sec
25	25.6	14.2	4800	2	US-08-612-734B-3	Sec

Case	Age	Sex	Ref.
1	27	C	US-09-134-001C-1460
2	25.4		US-09-134-001C-1460
3	14.1		US-09-134-001C-1460
4	14.2		US-09-134-001C-1460
5	17.0		US-09-134-001C-1460
6	14.2		US-09-134-001C-1460
7	14.2		US-09-134-001C-1460
8	14.2		US-09-134-001C-1460
9	14.2		US-09-134-001C-1460
10	14.2		US-09-134-001C-1460
11	14.2		US-09-134-001C-1460
12	14.2		US-09-134-001C-1460
13	14.2		US-09-134-001C-1460
14	14.2		US-09-134-001C-1460
15	14.2		US-09-134-001C-1460
16	14.2		US-09-134-001C-1460
17	14.2		US-09-134-001C-1460
18	14.2		US-09-134-001C-1460
19	14.2		US-09-134-001C-1460
20	14.2		US-09-134-001C-1460
21	14.2		US-09-134-001C-1460
22	14.2		US-09-134-001C-1460
23	14.2		US-09-134-001C-1460
24	14.2		US-09-134-001C-1460
25	14.2		US-09-134-001C-1460
26	14.2		US-09-134-001C-1460
27	14.2		US-09-134-001C-1460
28	14.2		US-09-134-001C-1460
29	14.2		US-09-134-001C-1460
30	14.2		US-09-134-001C-1460
31	14.2		US-09-134-001C-1460
32	14.2		US-09-134-001C-1460
33	14.2		US-09-134-001C-1460
34	14.2		US-09-134-001C-1460
35	14.2		US-09-134-001C-1460
36	14.2		US-09-134-001C-1460
37	14.2		US-09-134-001C-1460
38	14.2		US-09-134-001C-1460
39	14.2		US-09-134-001C-1460
40	14.2		US-09-134-001C-1460
41	14.2		US-09-134-001C-1460
42	14.2		US-09-134-001C-1460
43	14.2		US-09-134-001C-1460
44	14.2		US-09-134-001C-1460
45	14.2		US-09-134-001C-1460
46	14.2		US-09-134-001C-1460
47	14.2		US-09-134-001C-1460
48	14.2		US-09-134-001C-1460
49	14.2		US-09-134-001C-1460
50	14.2		US-09-134-001C-1460
51	14.2		US-09-134-001C-1460
52	14.2		US-09-134-001C-1460
53	14.2		US-09-134-001C-1460
54	14.2		US-09-134-001C-1460
55	14.2		US-09-134-001C-1460
56	14.2		US-09-134-001C-1460
57	14.2		US-09-134-001C-1460
58	14.2		US-09-134-001C-1460
59	14.2		US-09-134-001C-1460
60	14.2		US-09-134-001C-1460
61	14.2		US-09-134-001C-1460
62	14.2		US-09-134-001C-1460
63	14.2		US-09-134-001C-1460
64	14.2		US-09-134-001C-1460
65	14.2		US-09-134-001C-1460
66	14.2		US-09-134-001C-1460
67	14.2		US-09-134-001C-1460
68	14.2		US-09-134-001C-1460
69	14.2		US-09-134-001C-1460
70	14.2		US-09-134-001C-1460
71	14.2		US-09-134-001C-1460
72	14.2		US-09-134-001C-1460
73	14.2		US-09-134-001C-1460
74	14.2		US-09-134-001C-1460
75	14.2		US-09-134-001C-1460


```

: GENERAL INFORMATION:
: APPLICANT: Braun, Carl J.
: APPLICANT: Hemenway, Cynthia L.
: APPLICANT: Turner, Nilgun E.
: TITLE OF INVENTION: Plants Resistant to Infection by PVX
: NUMBER OF SEQUENCES: 12
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Dennis R. Hoerner, Jr., Monsanto Co. BB4F
: STREET: 700 Chesterfield Village Parkway
: CITY: St. Louis
: STATE: Missouri
: COUNTRY: USA
: ZIP: 63198
: COMPUTER READABLE FORM: disk
: MEDIUM TYPE: Floppy disk

```


Db 1823 GTCTCAACGGCTCTTTTC 1804

RESULT 13

US-08-803-972-1/c

; Sequence 1, Application US/08803972

; Patent No. 5792937

; GENERAL INFORMATION:

; APPLICANT: Braun, Carl J.

; APPLICANT: Hemenway, Cynthia L.

; APPLICANT: Turner, Nilgun E.

; TITLE OF INVENTION: Plants Resistant to Infection by PVX

; NUMBER OF SEQUENCES: 12

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Dennis R. Hoerner, Jr., Monsanto Co. BB4F

; STREET: 700 Chesterfield Village Parkway

; CITY: St. Louis

; STATE: Missouri

; COUNTRY: USA

; ZIP: 63198

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/803,972

; FILING DATE: 21-FEB-1997

; CLASSIFICATION: 800

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 07/804,862

; FILING DATE: 04-DEC-1991

; APPLICATION NUMBER: US 07/771,912

; FILING DATE: 04-OCT-1991

; ATTORNEY/AGENT INFORMATION:

; NAME: Hoerner Jr., Dennis R.

; REGISTRATION NUMBER: 30,914

; REFERENCE/DOCKET NUMBER: 38-21(10541)A

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (314)537-6099

; TELEFAX: (314)537-6047

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 4371 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

; MOLECULE TYPE: cDNA

; FEATURE:

; NAME/KEY: CDS

; LOCATION: 1..4371

US-08-803-972-1

Query Match 16.0%; Score 28.8; DB 1; Length 4371;

Best Local Similarity 60.0%; Pred. No. 1.6;

Matches 48; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

QY 27 GCAGTTGTTGACGAACCTCTTTTACAAACCGAATCGTTCTGAGTCATCTTTTGCCGCGC 86

Db 1883 GCATTTAGTTAGGAATCCAGTGTTCACAGGAGTCCTCTCTATGACTGTGTCGCTGCT 1824

QY 87 GAATCTATGGGAACCTTAC 106

Db 1823 GTCTCAACGGCTCTTTTC 1804

RESULT 14

US-08-781-891-79

; Sequence 79, Application US/08781891

; Patent No. 6090620

; GENERAL INFORMATION:

; APPLICANT: Fu, Ying-Hui

; APPLICANT: Yu, Chang-En
; APPLICANT: Oshima, Junko
; APPLICANT: Mulligan, John T.
; APPLICANT: Schellenberg, Gerald D.

; TITLE OF INVENTION: GENE AND GENE PRODUCTS RELATED TO

; NUMBER OF SEQUENCES: 209

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: SEED and BERRY LLP

; STREET: 6300 Columbia Center, 701 Fifth Avenue

; CITY: Seattle

; STATE: Washington

; COUNTRY: USA

; ZIP: 98104-7092

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/781,891

; FILING DATE: 27-DEC-1996

; CLASSIFICATION: 800

; ATTORNEY/AGENT INFORMATION:

; NAME: No. 6090620tenburg Ph.D., Carol

; REGISTRATION NUMBER: 39,317

; REFERENCE/DOCKET NUMBER: 240052.419

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (206) 622-4900

; TELEFAX: (206) 682-6031

; INFORMATION FOR SEQ ID NO: 79:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 87350 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

US-08-781-891-79

Query Match 15.7%; Score 28.2; DB 3; Length 87350;

Best Local Similarity 54.3%; Pred. No. 8.8;

Matches 57; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY 11 ATAATGCTTATCTGATGCGAGTTGTGTGACGAACCTCTTTTACACCGAATCGTTCTGA 70

Db 69319 ATAATACGTTTATGTGTACCTTGTATTATTATTATGATAGTAATAATTATATAC 69378

QY 71 GTCATCTTTTGGCGGAATCTATGGAACCTTACCTCATTTTC 115

Db 69379 ACACACGATATGCTCTAATCACTGGAGCCCTGCTGTTTTC 69423

RESULT 15

US-09-791-211-3

; Sequence 3, Application US/09791211

; Patent No. 6448080

; GENERAL INFORMATION:

; APPLICANT: Donna T. Ward

; APPLICANT: Andrew T. Watt

; TITLE OF INVENTION: ANTISENSE MODULATION OF WRN EXPRESSION

; FILE REFERENCE: RTS-0205

; CURRENT APPLICATION NUMBER: US/09/791,211

; CURRENT FILING DATE: 2001-02-23

; NUMBER OF SEQ ID NOS: 90

; SEQ ID NO 3

; LENGTH: 87543

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: unsure

; LOCATION: 7421

; OTHER INFORMATION: unknown

; NAME/KEY: unsure

; LOCATION: 7427

NAME/KEY:	unknown
LOCATION:	47291
OTHER INFORMATION:	unknown
NAME/KEY:	unknown
LOCATION:	52786
OTHER INFORMATION:	unknown
NAME/KEY:	unknown
LOCATION:	52787
OTHER INFORMATION:	unknown
NAME/KEY:	unknown
LOCATION:	53384
OTHER INFORMATION:	unknown
NAME/KEY:	unknown
LOCATION:	54684
OTHER INFORMATION:	unknown
NAME/KEY:	unknown
LOCATION:	59215
OTHER INFORMATION:	unknown
NAME/KEY:	unknown
LOCATION:	59235
OTHER INFORMATION:	unknown
NAME/KEY:	unknown
LOCATION:	59242
OTHER INFORMATION:	unknown
NAME/KEY:	unknown
LOCATION:	63290
OTHER INFORMATION:	unknown
NAME/KEY:	unknown
LOCATION:	66614
OTHER INFORMATION:	unknown
NAME/KEY:	unknown
LOCATION:	68660
OTHER INFORMATION:	unknown
NAME/KEY:	unknown
LOCATION:	68697
OTHER INFORMATION:	unknown
NAME/KEY:	unknown
LOCATION:	68718
OTHER INFORMATION:	unknown
NAME/KEY:	unknown
LOCATION:	68733
OTHER INFORMATION:	unknown
NAME/KEY:	unknown
LOCATION:	68739
OTHER INFORMATION:	unknown
NAME/KEY:	unknown
LOCATION:	69785
OTHER INFORMATION:	unknown
NAME/KEY:	unknown
LOCATION:	79134
OTHER INFORMATION:	unknown
NAME/KEY:	unknown
LOCATION:	79198
OTHER INFORMATION:	unknown
NAME/KEY:	unknown
LOCATION:	86336
OTHER INFORMATION:	unknown
OTHER INFORMATION:	US-09-791-211-3

Qy	11	ATAATGCTATCTGATGCGAGTTGTTGACGAAACCTCTTTTACAAACGGAATCGTTCGA	70
Db	69512	ATAATACGTTATTTGTGTACCTTGTTATATTATTCATGTAGTATAGAAATATATAC	69571
Qy	71	GTCACTCTTTGGCCGGGAATCATGCGAAACCTTACCTCATTTTC	115
Db	69572	ACACACGATAATGTCTAATAGTGGAGCCCTGAGCTGTGTTTC	69616

Search completed: June 6, 2003, 16:28:03
Job time : 37.9245 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 6, 2003, 10:02:20 ; Search time 423 Seconds
(without alignments)
7506.661 Million cell updates/sec

Title: US-09-775-879-20_COPY_143_1552

Perfect score: 1410

Sequence: 1 at:gcgcgcgataatgctta.....tgcatcatcggaacatctg 1410

Scoring table: IDENTITY_NUC

Gap: 10.0 , Gapext 1.0

Searched: 2185339 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_101002.*

1: /SID22/gcgcgata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SID22/gcgcgata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SID22/gcgcgata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SID22/gcgcgata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SID22/gcgcgata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SID22/gcgcgata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SID22/gcgcgata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SID22/gcgcgata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SID22/gcgcgata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SID22/gcgcgata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SID22/gcgcgata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SID22/gcgcgata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SID22/gcgcgata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SID22/gcgcgata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SID22/gcgcgata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SID22/gcgcgata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SID22/gcgcgata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SID22/gcgcgata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SID22/gcgcgata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SID22/gcgcgata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SID22/gcgcgata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SID22/gcgcgata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SID22/gcgcgata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SID22/gcgcgata/geneseq/geneseqn-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Quer/Match	Length	ID	Description
1	278.4	19.7	1164	21 AAC41897	Arabidopsis thaliana
2	164.2	11.5	1381	22 AAH50964	Sterol C5 desaturase
3	158.2	11.2	446	21 AAC69689	Eucalyptus grandis
4	125.4	8.3	285	24 ABL72194	Corn tassell-derive
5	106.2	7.5	304	24 ABL70846	Corn tassell-derive
6	54.8	3.3	6485	22 AAS46559	Tumour suppressor
7	54.8	3.3	6485	24 ABL33808	Human immune syste
8	54.2	3.3	6015	24 ABL32676	Human immune syste
9	54	3.3	7008	22 AAS46531	Tumour suppressor

c	10	53.6	3.8	875	22	AAI95044	Human neuroblastom
	11	52	3.7	9539	22	AAS45347	Chemically pretrea
	12	52	3.7	9539	24	ABK38180	DNA transcription
	13	51.4	3.6	6145	24	ABL32972	Human immune syste
	14	51.2	3.6	5181	24	ABL70443	Chemically treated
	15	51.2	3.6	5493	24	AAS61386	Human gene regulat
	16	51.2	3.6	12007	24	ABL32717	Human immune syste
c	17	51	3.6	693	24	ABQ32104	Oligonucleotide fo
	18	51	3.6	693	24	ABQ32105	Oligonucleotide fo
	19	50.8	3.6	6301	24	ABL32053	Human immune syste
	20	50.4	3.6	556	24	ABQ36996	Oligonucleotide fo
c	21	50.4	3.6	556	24	ABQ36997	Oligonucleotide fo
	22	50	3.5	5432	24	ABN80000	Human chemically m
c	23	49.8	3.5	6644	20	AAX33181	Base sequence of t
	24	49.8	3.5	7118	22	AAS46381	Tumour suppressor
	25	49.8	3.5	7118	24	ABL32830	Human immune syste
	26	49.8	3.5	7348	22	AAS46335	Tumour suppressor
c	27	49.8	3.5	7372	20	AAX33182	Base sequence of t
	28	49.8	3.5	7797	20	AAX33180	Cowpox virus bsr f
c	29	49.8	3.5	7996	20	AAX33184	Base sequence of t
	30	49.4	3.5	5269	24	ABL34056	Human immune syste
	31	49.4	3.5	12142	24	ABL33673	Human immune syste
	32	49.4	3.5	14006	24	ABL33958	Human immune syste
	33	49	3.5	47108	24	ABK31510	Signal transductio
	34	48.8	3.5	5586	24	ABK40004	Human chemically p
	35	48.8	3.5	6334	24	ABL33213	Human immune syste
	36	48.8	3.5	6668	24	ABL33697	Human immune syste
	37	48.8	3.5	17594	24	ABL34026	Human immune syste
	38	48.8	3.5	17934	24	ABL33719	Human immune syste
	39	48.4	3.4	5437	24	ABL33412	Human immune syste
	40	48.4	3.4	12465	22	AAS45326	Chemically pretrea
	41	48.4	3.4	12465	24	ABL32634	Human immune syste
c	42	48.2	3.4	819	22	AAI94974	Human neuroblastom
	43	48.2	3.4	6656	24	ABL33400	Human immune syste
c	44	48	3.4	626	23	ABV60941	Human prostate exp.
	45	48	3.4	6794	24	ABL70220	Chemically treated

ALIGNMENTS

RESULT 1

AAC41897
ID AAC41897 standard; DNA; 1164 BP.
XX AAC41897;
AC AAC41897;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 33540.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131149.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138340.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.

PT cancer -
PS Claim 1; SEQ ID NO 281; 27pp; English.
XX
XX The invention relates to a nucleic acid comprising a sequence of 18
CC bases, of a segment of chemically pretreated DNA (CP DNA) e.g. with
CC bisulphite, of genes associated with tumour suppression and
CC oncogenes having a sequence taken from 536 (actually 533 since
CC numbers 408, 458 and 500 are missing from the sequence listing) sequences
CC (SS) and sequences complementary to (SS). The nucleic acid may be a
CC peptide nucleic acid-oligomer (PNA) of at least 9 nucleotides and may
CC form part of a set of probes for detecting the cytosine methylation state
CC and/or single nucleotide polymorphisms and also to be used in an
CC array for analysing diseases associated with CpG dinucleotides e.g.
CC cancers and tumours. The probes can also be used in a method for
CC ascertaining genetic and/or epigenetic parameters for the diagnosis
CC and/or therapy of existing diseases or the predisposition to specific
CC diseases, by analysing cytosine methylations. The parameters may be
CC compared to another set of genetic and/or epigenetic parameters, the
CC differences serving as basis for diagnosis and/or prognosis events which
CC are disadvantageous to patients. The present sequence is one of the
CC 533 genomic sequences derived from tumour suppressor genes and
CC oncogenes.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC fcp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 6485 BP; 1509 A; 168 C; 1822 G; 2985 T; 1 other;

Query Match 3.9%; Score 54.8; DB 22; Length 6485;
Best Local Similarity 46.2%; Pred. No. 0.00068;
Matches 182; Conservative 0; Mismatches 212; Indels 0; Gaps 0;
QY 191 TCTATTACCTTAAATCAACGTTTACCTTCCCAAGGTCGACATTCACATTTGTATTC 250
Db 2363 TTTTATAGTTTAAAGTTTATTCGTTTGTGATATGTTTTTCGTTTTTTTTTTTTT 2422
QY 251 ACTATTGCTTAATCGCTTCTATGTATCGATTTTCAATTTAAGGAAGAGGTTTCTTC 310
Db 2423 TTTTATAGTTTAAAGTTTATTCGTTTGTGATATGTTTTTCGTTTTTTTTTTTTT 2482
QY 311 GTGTAAGTAAATTTGGATTTGATGAGATTCATGTTCCATTTATGATAT 370
Db 2483 ATTTTGTTTTTTTTTTTTTTTTTTTTTTTTTTTTTCGTTTATTTTTCGTTTTT 2542
QY 371 TTGTGCATATCTCCATCTAAGGATTAACAGTTAGTGGCTATATATAAGTTTCTGCA 430
Db 2543 TTTTATAGTTTAAAGTTTATTCGTTTGTGATATGTTTTTCGTTTTTTTTTTTTT 2602
QY 431 ACCAATGAGAAGTCGTACATCTTTGAAGTTTCAATTTTCTACTTGCACATTAAGTCCACT 490
Db 2603 TTTTATAGTTTAAAGTTTATTCGTTTGTGATATGTTTTTCGTTTTTTTTTTTTT 2662
QY 491 AAATGTTTGTGAAGTGAATGCTACTTTCAGACACATCTTTTCTGCTCTCGAGA 550
Db 2663 TTAATTTTTTTTTTTTTTTTTTTTTTTTTTTAGGTTTATTTTATTTTTTTTTTTTTT 2722
QY 551 CTCGTCTTAGTTTGAATCTTTTTGGTCTGTT 584
Db 2723 TTTTATAGTTTAAAGTTTATTCGTTTGTGATATGTTTTTCGTTTTTTTTTTTTT 2756

RESULT 7

ABL33808
ID ABL33808 standard; DNA; 6485 BP.
XX
AC ABL33808;
XX
DT 26-MAR-2002 (first entry)
XX
DE Human immune system associated gene SEQ ID NO: 1781.
XX

KW Human; immune system disease; cytosine methylation; antiasthmatic;
KW antiarteriosclerotic; antianaemic; cytosine; cytosine; cytosine;
KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
KW antineuritic; antiarthritic; antidiabetic; antipsoriatic;
KW antineuritic; cancer; eye disease; arteriosclerosis; anaemia;
KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;
KW gene; ds.
XX Homo sapiens.
XX WO200200928-A2.
PN 03-JAN-2002.
XX 02-JUL-2001; 2001WO-EP07537.
XX 30-JUN-2000; 2000DE-1032529.
PR 01-SEP-2000; 2000DE-1043826.
XX (EPIG-) EPIGENOMICS AG.
XX Olek A, Piepenbrock C, Berlin K;
PI WPI; 2002-130909/17.
XX Nucleic acid comprising fragment of chemically modified gene, useful
PT for diagnosis and treatment of diseases associated with abnormal
PT cytosine methylation -
XX Claim 1; SEQ ID NO 1781; 32pp + Sequence Listing; German.
PS The present invention provides a number of human immune system associated
XX genes which are modified by the methylation of cytosines. The sequences
CC can be used in the diagnosis and treatment of immune system disorders,
CC including eye diseases such as retinopathy, neovascular glaucoma and
CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
CC diseases. The present sequence is a gene of the invention.
XX
XX Sequence 6485 BP; 1509 A; 168 C; 1822 G; 2985 T; 1 other;

Query Match 3.9%; Score 54.8; DB 24; Length 6485;
Best Local Similarity 46.2%; Pred. No. 0.00068;
Matches 182; Conservative 0; Mismatches 212; Indels 0; Gaps 0;
QY 191 TCTATTACCTTAAATCAACGTTTACCTTCCCAAGGTCGACATTCACATTTGTATTC 250
Db 2363 TTTTATAGTTTAAAGTTTATTCGTTTGTGATATGTTTTTCGTTTTTTTTTTTTT 2422
QY 251 ACTATTGCTTAATCGCTTCTATGTATCGATTTTCAATTTAAGGAAGAGGTTTCTTC 310
Db 2423 TTTTATAGTTTAAAGTTTATTCGTTTGTGATATGTTTTTCGTTTTTTTTTTTTT 2482
QY 311 GTGTAAGTAAATTTGGATTTGATGAGATTCATGTTTCGATTTTATGATAT 370
Db 2483 ATTTTGTTTTTTTTTTTTTTTTTTTTTTTTTTTAGGTTTATTTTATTTTTTTTTT 2542
QY 371 TTGTGCATATCTCCATCTAAGGATTAACAGTTAGTGGCTATATATAAGTTTCTGCA 430
Db 2543 TTTTATAGTTTAAAGTTTATTCGTTTGTGATATGTTTTTCGTTTTTTTTTTTTT 2602
QY 431 ACCAATGAGAAGTCGTACATCTTTGAAGTTTCAATTTTCTACTTGCACATTAAGTCCACT 490
Db 2603 TTTTATAGTTTAAAGTTTATTCGTTTGTGATATGTTTTTCGTTTTTTTTTTTTT 2662
QY 491 AAATGTTTGTGAAGTGAATGCTACTTTCAGACACATCTTTTCTGCTCTCGAGA 550
Db 2663 TTAATTTTTTTTTTTTTTTTTTTTTTTTTTTAGGTTTATTTTATTTTTTTTTTTTTT 2722
QY 551 CTCGTCTTAGTTTGAATCTTTTTGGTCTGTT 584

```
Db 2723 TTTTTFMTTTTTTTTTTTTTTTTATTATTATTCGTGTT 2756
RESULT 8
ABL32676
ID ABL32676 standard; DNA; 6015 BP.
XX
AC ABL32676;
XX
DT 26-MAR-2002 (first entry)
XX
DE Human immune system associated gene SEQ ID NO: 649.
XX
KW Human; immune system disease; cytosine methylation; antiasthmatic;
KW antiarteriosclerotic; antianaemic; cytostatic; nootropic;
KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
KW antirheumatic; antiarthritic; antidiabetic; antipsoriatic;
KW antinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;
KW gene; ds.
XX
OS Homo sapiens.
XX
PN WO200200928-A2.
XX
PD 03-JAN-2002.
XX
PF 02-JUL-2001; 2001WO-EP07537.
XX
PR 30-JUN-2000; 2000DE-1032529.
XX
PR 01-SEP-2000; 2000DE-1043826.
XX
PA (EPIG-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K;
XX
DR WPI; 2002-130909/17.
XX
PT Nucleic acid comprising fragment of chemically modified gene, useful
PT for diagnosis and treatment of diseases associated with abnormal
PT cytosine methylation
XX
PS Claim 1; SEQ ID NO 649; 32pp + Sequence Listing; German.
XX
CC The present invention provides a number of human immune system associated
CC genes which are modified by the methylation of cytosines. The sequences
CC can be used in the diagnosis and treatment of immune system disorders,
CC including eye diseases such as retinopathy, neovascular glaucoma and
CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
CC diseases. The present sequence is a gene of the invention.
XX
SQ Sequence 6015 EP; 1822 A; 96 C; 1277 G; 2820 T; 0 other;
Query Match 3.8%; Score 54.2; DB 24; Length 6015;
Best Local Similarity 46.6%; Pred No. 0.00095;
Matches 173; Conservative 0; Mismatches 198; Indels 0; Gaps 0;
QY 243 TTGTAITACATTCGTTAAATCGCTTCTATGTTATCGATTTCATTTAAGGAGAGG 302
Db 1012 TTGTCGTTTGTGTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTG 1071
QY 303 GTTCTTCGTACTGTACAGTAATTCGATTGATGAGTGGATGATTCATGTTTCGATTTA 362
Db 1072 TGGGGTATGTGAAATTCGTGTATGTATAGATGCTGTAATGATTAGTAGGATGCTT 1131
QY 363 TTGATATTGTCATATTCCTCACTAAGGATTTGAACAGTTAGTCGCTTATATACTT 422
Db 1132 GGTATTTTATTTATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 1191
QY 423 TTGTGCAACCAATGAGAAGTCGTACATCTTTGAAGTTGAATTTTCTACTTGCCATTTAA 482
Db 1192 ATTTAGAAAATTATAATATTTTGTGTTTATTGTAGTTATTTATTTTGTGTTAATAT 1251
QY 483 GTCCACTTAAATTTGTTTGAAGTATTCTACTTTTCACACACATTCCTTTTCTGCTT 542
Db 1252 TAGAAAGTATTTTATTTTAAATTTTGTGTTGTTATTTTATTAATTTTCTGTT 1311
QY 543 CTCTGAGACTCTGCTTAAATCTTTTGTGTTGTTTCTGTTTCTGTTTCTGTTTCTGTT 602
Db 1312 TTTATTTTATTTTATATATATATTTTATTTTATTTTATTTTATTTATTTTATTT 1371
QY 603 CTACAATAAAG 613
Db 1372 TTACGGTTATG 1382
RESULT 9
AAS46531
ID AAS46531 standard; DNA; 7008 BP.
XX
AC AAS46531;
XX
DT 18-DEC-2001 (first entry)
XX
DE Tumour suppressor gene derived chemically modified sequence #253.
XX
KW Human; tumour suppressor gene; oncogene; antitumour; cytostatic;
KW cancer; tumour; CpG dinucleotide; single-nucleotide polymorphism; SNP;
KW cytosine methylation; ds.
XX
OS Homo sapiens.
XX
PN WO200168912-A2.
XX
PD 20-SEP-2001.
XX
PF 15-MAR-2001; 2001WO-EP02955.
XX
PR 15-MAR-2000; 2000DE-1013847.
XX
PR 06-APR-2000; 2000DE-1019058.
XX
PR 07-APR-2000; 2000DE-1019173.
XX
PR 30-JUN-2000; 2000DE-1032529.
XX
PR 01-SEP-2000; 2000DE-1043826.
XX
PA (EPIG-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K;
XX
DR WPI; 2001-602752/68.
XX
PT Fragments of chemically modified genes associated with tumour suppressor
PT genes and oncogenes, useful in designing primers and probes for
PT analysing diseases associated with cytosine methylation state e.g.
PT cancer
XX
PS Claim 1; SEQ ID No 253; 27pp; English.
XX
CC The invention relates to a nucleic acid comprising a sequence of 18
CC bases, of a segment of chemically pretreated DNA (CP DNA) e.g. with
CC bisulphite, of genes associated with tumour suppression and
CC oncogenes having a sequence taken from 536 (actually 533 since
CC numbers 408, 458 and 500 are missing from the sequence listing) sequences
CC (SS) and sequences complementary to (SS). The nucleic acid may be a
CC peptide nucleic acid-oligomer (PNA) of at least 9 nucleotides and may
CC form part of a set of probes for detecting the cytosine methylation state
CC and/or single nucleotide polymorphisms and also to be used in an
CC array for analysing diseases associated with CpG dinucleotides e.g.
CC cancers and tumours. The probes can also be used in a method for
CC ascertaining genetic and/or epigenetic parameters for the diagnosis
CC and/or therapy of existing diseases or the predisposition to specific
CC diseases, by analysing cytosine methylations. The parameters may be
CC compared to another set of genetic and/or epigenetic parameters, the
CC differences serving as basis for diagnosis and/or prognosis events which
```


THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 6, 2003, 13:35:44 ; Search time 250.075 Seconds
(without alignments)
1729.135 Million cell updates/sec

Title: US-09-775-879-20_COPY_143_1552
Perfect score: 1410
Sequence: 1 atgcggcgataatgctta.....tgcatccatgcgaacatctg 1410

Scoring table: IDENT:TY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 44136; seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA.*
1: /3gn2_6/ptodata/1/1na/5A_COMB.seq.*
2: /3gn2_6/ptodata/1/1na/5B_COMB.seq.*
3: /3gn2_6/ptodata/1/1na/6A_COMB.seq.*
4: /3gn2_6/ptodata/1/1na/6B_COMB.seq.*
5: /3gn2_6/ptodata/1/1na/PCTUS_COMB.seq.*
6: /3gn2_6/ptodata/1/1na/backfiles.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	184	13.0	1447	4	US-09-443-041A-27
2	166	11.8	911	4	US-09-443-041A-23
3	165.8	11.8	1070	4	US-09-443-041A-9
4	163.4	11.6	1210	4	US-09-443-041A-29
5	161.8	11.5	1318	4	US-09-443-041A-25
6	92.8	6.6	495	4	US-09-443-041A-15
7	89	6.3	459	4	US-09-443-041A-13
8	65	4.6	360	4	US-09-443-041A-11
9	55.6	3.5	7218	1	US-08-232-463-14
c 10	42.8	3.0	5058	4	US-09-889-595-1
c 11	39.6	2.8	19124	2	US-08-487-826B-13
c 12	38.2	2.7	43676	3	US-09-356-952-12
c 13	37.2	2.6	2060	1	US-07-721-761A-31
c 14	37.2	2.6	2060	1	US-07-978-687-31
c 15	37.2	2.6	2060	5	PCT-US91-01750-2
c 16	37.2	2.6	2060	5	PCT-US91-05801-31
c 17	37.2	2.6	2081	5	PCT-US91-01750-3
c 18	37.2	2.6	3440	1	US-08-471-791-27
c 19	37.2	2.6	3440	5	PCT-US91-01746-27
20	37	2.6	4817	1	US-07-951-715A-18
21	37	2.6	4817	3	US-08-459-448A-18
22	37	2.6	4817	3	US-08-459-595A-18
23	37	2.6	4817	3	US-08-459-504B-18
24	37	2.6	4817	3	US-08-459-444-18
25	37	2.6	4817	4	US-09-547-422-18
26	36.4	2.6	658	4	US-08-998-416-595
27	36.2	2.6	731	4	US-08-887-534A-8

28	36	2.6	573	4	US-08-936-165A-48	Sequence 48, Appl
29	35.8	2.5	1622	4	US-08-949-246-3	Sequence 3, Appl
c 30	35.6	2.5	2484	4	US-08-961-527-234	Sequence 234, App
31	35.4	2.5	10968	2	US-08-680-327-2	Sequence 2, Appl
32	35.4	2.5	10968	4	US-09-228-246-1	Sequence 1, Appl
c 33	35.2	2.5	289	4	US-09-007-005-17	Sequence 17, Appl
c 34	35.2	2.5	289	4	US-09-244-796-17	Sequence 17, Appl
c 35	34.8	2.5	41708	4	US-09-470-512A-3	Sequence 3, Appl
36	34.8	2.5	51259	3	US-08-781-891-209	Sequence 209, App
c 37	34.6	2.5	327	4	US-09-889-595-7	Sequence 7, Appl
c 38	34.6	2.5	3457	1	US-08-295-882-1	Sequence 1, Appl
c 39	34.2	2.4	6243	2	US-09-056-075-1	Sequence 1, Appl
c 40	34	2.4	51259	3	US-08-781-891-209	Sequence 209, App
41	34	2.4	112132	4	US-09-741-150-3	Sequence 3, Appl
c 42	33.8	2.4	2370	1	US-08-104-072B-7	Sequence 7, Appl
c 43	33.8	2.4	2370	1	US-08-351-413-8	Sequence 8, Appl
c 44	33.8	2.4	2370	2	US-09-025-583-8	Sequence 8, Appl
c 45	33.8	2.4	9636	1	US-08-323-170B-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-09-443-041A-27
; Sequence 27, Application US/09443041A
; Patent No. 6465717
; GENERAL INFORMATION:
; APPLICANT: Famodu, Omolayo O.
; APPLICANT: Orozco, Buddy
; APPLICANT: Rafalski, Antoni
; APPLICANT: Shen, Jennie
; TITLE OF INVENTION: Sterol Metabolism Enzymes
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/09/443,041A
; CURRENT FILING DATE: 1999-11-18
; PRIOR APPLICATION NUMBER: 60/109,283
; PRIOR FILING DATE: 1998-11-20
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 27
; LENGTH: 1447
; TYPE: DNA
; ORGANISM: Glycine max
US-09-443-041A-27

Query Match	13.0%	Score	184;	DB	4;	Length	1447;
Best Local Similarity	72.6%	Pred. No.	3.8e-42;				
Matches	238;	Conservative	0;	Mismatches	90;	Indels	0;
Gaps	0;						
Qy	592	AGATGCAATTCCTACAAATAAAGGCTATGCGTTTGGCAAAATGTTGTGGCAATGAAGGCTAT	651				
Db	316	AGATGCTATTCCTCTCAAGAGCCATGCTCTTGCNAATATCTGTTGTATGAAGCCAT	375				
Qy	652	GCCATGTPACACTCTTCTTCCAACTGTCCTCGAGAGTAGTATGATTAACCTGTTGGACCAA	711				
Db	376	GCCGTTGTTACTCTTTTACTTCCAACTGTTTTCGGAGTACCTGTTAGAACTGGTGGACAAA	435				
Qy	712	ATGTTTCTAGCATAGACGAATTCGGCTGATCTGTATTTTGTATTCATCGCCATCTA	771				
Db	436	GTGTCATCTCTATATATAATGTTGGTTGTCATACCTTGTATTTAGCAATTTA	495				
Qy	772	TCTTGTTCCTGAGTTTGGTATTTATTTGATGACAGAGCTTTCATGACATTAAGCC	831				
Db	496	TCTATATTTAGAGTTTGGTATTTATTTGATGACAGAGCTTTCATGACATTAAGCC	555				
Qy	832	TCTTATATGATTCCTCCATGCCACCCATCATATATACAAAGCAAGATACATCTCTCC	891				
Db	556	GCTTTACAAATATCTTCATGCTACCCATCATCTACAAATAAAGCAAGCAACTCTCTCC	615				
Qy	892	ATTTGCCGTAAGTGTTCATGTTTTCAGTTTGT	919				
Db	616	TTTGTGTTGGCATTTTCACCTCTT	643				

; OTHER INFORMATION: any nucleotide
US-09-443-041A-29

Query Match 11.6%; Score 163.4; DB 4; Length 1210;
Best Local Similarity 70.6%; Pred. No. 2.1e-36;
Matches 218; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

QY 592 AGATGCAATTCCTACAAATAAGGCTATGCGTTTGGCAATGTTGGCAATGAAGGCTAT 651
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 188 AGATGCTATCCCTACAGTAGAGCTATGAAGAAGCAATAATTTGGCATCAAGGCTAT 247
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

QY 652 GCCATGGTACACTCTCTCCAACTGCTCCGAGAGTAGATGAACGTGGTTGGACCAA 711
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 248 GCCCTTCCTACTGCTCTCCGCTGATCTGAGCAGCATGATTGAGAGTGGATGGACAG 307
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

QY 712 ATGTTTTCTAGCATAGAGCAATTCGGCTGGATCTGTTATTTGTTTACATCGCCATCTA 771
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Db 308 GTGTTCTTCATATACAGGAAGTTGGTGGCTATGACATATCTATGTCCTCTATA 367
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

QY 772 TCTGTTTCGTTGAGTTGGTATTATTGGATGCACAGAGCTTCATGACATTAAGCC 831
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Db 368 TCTCATCTTTGGGAGTTCCGAATTTACTGGATGCACAGAGTTGCGATGACATAAGCC 427
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

QY 832 TCTCTATAGTATCTCCATGCCACCCATCATATCTACAAACAGAGAGATACACTCTCTCC 891
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Db 428 ACTATACAGACCTACATGTCACCCACCATTTTACAAACAGAGAGATACCCCTATCACC 487
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

QY 892 ATTTGCCCG 900
||||| ||

Db 488 ATTTGCTGG 496
||||| ||

RESULT 5

US-09-443-041A-25
; Sequence 25, Application US/09443041A
; Patent No. 6465717
; GENERAL INFORMATION:
; APPLICANT: Farnodu, Omolayo O.
; APPLICANT: Orozco, Buddy
; APPLICANT: Rafalski, Antoni
; APPLICANT: Shen, Jennie
; TITLE OF INVENTION: Sterol Metabolism Enzymes
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/09/443,041A
; CURRENT FILING DATE: 1999-11-18
; PRIOR APPLICATION NUMBER: 60/109,283
; PRIOR FILING DATE: 1998-11-20
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 25
; LENGTH: 1318
; TYPE: DNA
; ORGANISM: Oryza sativa
US-09-443-041A-25

Query Match 11.5%; Score 161.8; DB 4; Length 1318;
Best Local Similarity 70.2%; Pred. No. 6.2e-36;
Matches 217; Conservative 0; Mismatches 92; Indels 0; Gaps 0;

QY 592 AGATGCAATTCCTACAAATAAGGCTATGCGTTTGGCAATGTTGGCAATGAAGGCTAT 651
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 351 AGATGCTATCCCTACAGTAGAGCTATGAAGAAGCAATAATTTGGCATCAAGGCTAT 410
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

QY 652 GCCATGGTACACTCTCTCCAACTGCTCCGAGAGTAGATGAACGTGGTTGGACCAA 711
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Db 411 GCCTCTCTATTGTCCTTCCAACTTATCTGAGTACATGTTGGAATGATGAGACACA 470
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

QY 712 ATGTTTCTAGCATAGAGCAATTCGCTGGATTCGTATTTGTTTACATCGCCATCTA 771
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Db 471 GTGTTATGTTATATACAGTAGAGTTGGTTGGCAATGATACCTGGTTTATCTGGCTTTATA 530
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

QY 772 TCTGTTTTCGTTGAGTTTGGTATTATTATGATGCACAGAGCTTCATGACATTAAGCC 831
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Db 531 TCTTATCTTTGTTGAGTTTGGAAATTTACTGGATGCACAGAGTTGCGATGACATAAGCC 590
QY 832 TCTCTATAAGTATCTCCATGCCCACCATCATATCTACAAACAGAGAGATACACTCTCTCC 891
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Db 591 ATTGACAGTAGTACCTGCGACACATACCATCATATTTTACAAACAGAGAGATACCTATCACC 650
QY 892 ATTTGCCCG 900
||||| ||

Db 651 ATTTGCAGG 659
||||| ||

RESULT 6

US-09-443-041A-15
; Sequence 15, Application US/09443041A
; Patent No. 6465717
; GENERAL INFORMATION:
; APPLICANT: Farnodu, Omolayo O.
; APPLICANT: Orozco, Buddy
; APPLICANT: Rafalski, Antoni
; APPLICANT: Shen, Jennie
; TITLE OF INVENTION: Sterol Metabolism Enzymes
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/09/443,041A
; CURRENT FILING DATE: 1999-11-18
; PRIOR APPLICATION NUMBER: 60/109,283
; PRIOR FILING DATE: 1998-11-20
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 15
; LENGTH: 495
; TYPE: DNA
; ORGANISM: Triticum aestivum
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (261)
; OTHER INFORMATION: any nucleotide
; NAME/KEY: unsure
; LOCATION: (286)
; OTHER INFORMATION: any nucleotide
; NAME/KEY: unsure
; LOCATION: (288)
; OTHER INFORMATION: any nucleotide
; NAME/KEY: unsure
; LOCATION: (328)
; OTHER INFORMATION: any nucleotide
; NAME/KEY: unsure
; LOCATION: (344)
; OTHER INFORMATION: any nucleotide
; NAME/KEY: unsure
; LOCATION: (372)
; OTHER INFORMATION: any nucleotide
; NAME/KEY: unsure
; LOCATION: (390)..(391)
; OTHER INFORMATION: any nucleotide
; NAME/KEY: unsure
; LOCATION: (450)..(451)
; OTHER INFORMATION: any nucleotide
; NAME/KEY: unsure
; LOCATION: (480)
; OTHER INFORMATION: any nucleotide
; NAME/KEY: unsure
; LOCATION: (489)
; OTHER INFORMATION: any nucleotide
US-09-443-041A-15

Query Match 6.6%; Score 92.8; DB 4; Length 495;
Best Local Similarity 64.1%; Pred. No. 9.3e-17;
Matches 164; Conservative 0; Mismatches 90; Indels 2; Gaps 2;

QY 592 AGATGCAATTCCTACAAATAAGGCTATGCGTTTGGCAATGTTGGCAATGAAGGCTAT 651
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Db 181 AGATGCTGCTCCCTACAGTAGAGCTATGAAGAAGCAATAATTTGTCATCAAGGCTAT 240
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: pT2gpt-F1s
US-08-232-463-14

Query Match
Best Local Similarity 3.9%; Score 55.6; DB 1; Length 7218;
Matches 10; Conservative 227; Mismatches 151; Indels 0; Gaps 0;

QY 1 ATGCGGGGGAATGCTATCTGATGCGAGTTGTTGACGAAACCTCTTTTACACCGA 60
DB 1047 AGGTGCGAAGGAGCTTGGCAATTTTTTTTTTTTTTTTTTTTTTTTTTTTTT
QY 61 ATCGTTCAGCTATCTTTGCGCGCAATCTATGGGAACCTTACCTCATTTCTCCAG 120
DB 1107 YY
QY 121 ACATGGCPCGAATTAACCTCGCGGAACCTACTATCTATCTCCGGTTCTCTCGG 180
DB 1167 YY
QY 181 TGCCTCATCTATTAACCTTAAATCAACGTTTACCTCCCAAGGCTCGACTTTCCAC 240
DB 1227 YY
QY 241 TTTTGTATCACTATGCTTAATCGCTTCTATGATGATGATGATGATGATGATGAT 300
DB 1297 YY
QY 301 GGGTTTCTGCTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360
DB 1347 YY
QY 361 TATTGATATTTGTCATATTTCTCCATC 388
DB 1407 YY

RESULT 10

US-09-889-595-1/c

Sequence 1, Application US/09889595

Patent No. 6410749

GENERAL INFORMATION:

APPLICANT: Aventis CropScience GmbH

TITLE OF INVENTION: PROMOTERS FOR GENE EXPRESSION IN CARYOPSES OF PLANTS

FILE REFERENCE: 5.4413-3885

CURRENT APPLICATION NUMBER: US/09/889,595

CURRENT FILING DATE: 2001-07-05

PRIOR APPLICATION NUMBER: DE 100 32 379.0

PRIOR FILING DATE: 2000-07-05

NUMBER OF SEQ ID NOS: 18

SOFTWARE: PatentIn version 3.0

SEQ ID NO 1

LENGTH: 5058

TYPE: DNA

ORGANISM: Triticum aestivum

US-09-889-595-1

Query Match
Best Local Similarity 3.0%; Score 42.8; DB 4; Length 5058;
Matches 158; Conservative 0; Mismatches 192; Indels 0; Gaps 0;

QY 271 TATGTTATCGATTTTCAATTAAGGAAGAGGGTTTCTGCTGCTACTGATCAAGTAATTTG 330
DB 3311 TTTTAAATATTTTCAATGCTGATTAACATTTTGAATATAATATAATCAAGATTTTCT 3252
QY 331 GATTGATGCTGATGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTG 390
DB 3251 GGGTACATTTCTATACAAATTTAACTGTTTTTCTAAATGCTTGATTTTAAATACA 3192

QY 391 AGGATTGAACAGTTAGTGGGTTATATAAGTTTTTGTGCAACCAATGAGAAGTCGTACAT 450
DB 3191 AGATTAAACATTTTTTAATACATGGTCAACATTTTGTATACACGTTTAAATTTTCAAA 3132
QY 451 CTTTGAAGTTGAAATTTTCTACTTGCCATTTTAACTTAAATTTGTTTGTGTAAGTGTAT 510
DB 3131 TCCTTTATTAACATTTTTTAAATATTTGTTTTTAAATGCTTGTTATTTTTTAAATAATAC 3072
QY 511 TGTCTACTTTTCAGACACATTTCTTTTCTGCTTCTGAGACTCTCTCTTAGTTTGAATC 570
DB 3071 ATGATAATTTTATACAATTTTAAATACATGTTAACTTATACACGTTTAAACATTT 3012
QY 571 TTTTGTGCTGCTGTTTGTCTCAGATGCAATTCCTACAATAAGGCTATGC 620
DB 3011 TTTCTAATACCTGTTTATTTTCTTTCTTCAATGCTAGATTTAAATACTTGC 2962

RESULT 11

US-08-487-826B-13/c

Sequence 13, Application US/08487826B

Patent No. 5993827

GENERAL INFORMATION:

APPLICANT: Sim, Kim L.

APPLICANT: Chitnis, Chetan

APPLICANT: Miller, Louis H.

APPLICANT: Peterson, David S.

APPLICANT: Su, Xin-zhaun

APPLICANT: Wellem, Thomas E.

TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX

TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS

NUMBER OF SEQUENCES: 45

CORRESPONDENCE ADDRESS:

ADDRESS: Knobbe Martens Olson & Bear

STREET: 620 Newport Center Drive

CITY: Newport Beach

STATE: California

COUNTRY: US

ZIP: 92660

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/487,826B

FILING DATE: 10-SEP-1993

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Israel, Ned

REGISTRATION NUMBER: 29,655

REFERENCE/DOCKET NUMBER: NIH121.001CP1

TELECOMMUNICATION INFORMATION:

TELEPHONE: (619) 235-8550

TELEFAX: (619) 235-0176

INFORMATION FOR SEQ ID NO: 13:

SEQUENCE CHARACTERISTICS:

LENGTH: 19124 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

HYPOTHETICAL: NO

ANTI-SENSE: NO

US-08-487-826B-13

Query Match
Best Local Similarity 2.8%; Score 39.6; DB 2; Length 19124;
Matches 126; Conservative 0; Mismatches 144; Indels 0; Gaps 0;

QY 317 TGTACAGTAATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTG 376
DB 15933 TTTTATTTTTTTTTTTTATTTTCAATTAATTTTTTTTTTAAATTTTAAATTTT 15874

NUMBER OF SEQUENCES: 51
CORRESPONDENCE ADDRESS:
ADDRESS: Calgene, Inc.
STREET: 1920 Fifth Street
CITY: Davis
STATE: CA
COUNTRY: USA
ZIP: 95616
COMPUTER READABLE FORM:
MEDIUM TYPE: 5 1/4 inch, 3.50 inch, 1.0 MB
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Macintosh 6.0.7
SOFTWARE: Microsoft Word 4.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/978,687
FILING DATE: FEBRUARY 1, 1993
CLASSIFICATION: 800
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: PCT/US91/05801
FILING DATE: 15-AUGUST-1991
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 07/568,493
FILING DATE: 15-AUGUST-1990
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 07/721,761
FILING DATE: 26-JUNE-1991
ATTORNEY/AGENT INFORMATION:
NAME: Elizabeth Lassen
REGISTRATION NUMBER: 31,845
NAME: Donna E. Scherer
REGISTRATION NUMBER: 34,719
REFERENCE/DOCKET NUMBER: CGNE 76-2 WO
TELECOMMUNICATION INFORMATION:
TELEPHONE: (916) 753-6313
TELEFAX: (916) 753-1510
INFORMATION FOR SEQ ID NO: 31:
SEQUENCE CHARACTERISTICS:
LENGTH: 2060 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: genomic DNA
US-07-978-687-31

Query Match	2.6%	Score 37.2;	DB 1;	Length 2060;
Best Local Similarity	46.2%;	Pred. No. 0.73;		
Matches 123;	Conservative 0;	Mismatches 143;	Indels 0;	Gaps 0;
QY	365	GATTATTTGTGCAATATCTCCATCTCAAGGATTGAACAGTTAGTGGCTATATAAGTTTT	424	
Db	300	GATTAAAGTGGGAGATTCATTTTTTAAGGATTTGAAGAAATTGAAGATTCATTTGTTATT	241	
QY	425	TGTGCA/CCAANTCAGAGTCGTACACATTTTGAAGTTCAATTTTCTACTTCCCATTTAAGT	484	
Db	240	GCATTATTTTATCTTAAATTTTTTAAAGTCCTTGTGTATTGGTGTGTTGATTTTATAAT	181	
QY	485	CCACTTTAAATGTTTGTGGAAGTGATTGTCCTACTTTCAGACACATCTTTTTTCGCTCT	544	
Db	180	TCCATA/AAATCTTTTGTATTGCGTACACAATTTTCTAACTCAATTAATTTTTTTTATT	121	
QY	545	CTGAGA/CTCTGCTTATAGTTTGAATCTTTTTTTTGGTCTGTTTCTCAGATGCAATTCCT	604	
Db	120	ATTAAA/AAAGCTCAGTATTTTCGTGTTTTTTTTTTTAACTCTTGTACACAAATATTGT	61	
QY	605	ACAATAAAGCTATAGCTTTTGCAAT	630	
Db	60	TACACA/AAAAATGTCACCTTTTATAAT	35	

RESULT 15

RESULT IS
PCT-US91-01750-2/c

PC1-0391-01750-2/C ; Sequence 2, Application PC/TUS9101750

; sequence 2, Application:
; GENERAL INFORMATION:
; GENERAL INFORMATION:

APPLICANT: KNAUF, VIC C.
 APPLICANT: KRIDL, JEAN C.
 APPLICANT: SCHERER, DONNA E.
 TITLE OF INVENTION: Novel Sequences Preferentially
 Expresed In Early Seed
 TITLE OF INVENTION: Development and Methods
 Related Thereto
 NUMBER OF SEQUENCES: 13
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Calgene, Inc.
 STREET: 1920 Fifth Street
 CITY: Davis
 STATE: CA
 COUNTRY: USA
 ZIP: 95616
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette, 3.50 inch, 1.0 MB storage
 COMPUTER: Apple Macintosh
 OPERATING SYSTEM: Macintosh 6.0
 SOFTWARE: Microsoft Word 4.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: PCT/US91/01750
 FILING DATE: 19910314
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 07/494,722
 FILING DATE: 16-MAR-1990
 ATTORNEY/AGENT INFORMATION:
 NAME: Elizabeth Lassen
 REGISTRATION NUMBER: 31,845
 NAME: Donna E. Scherer
 REGISTRATION NUMBER: 34,719
 REFERENCE/DOCKET NUMBER: CGNE 68WO
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 916-753-6313
 TELEFAX: 916-753-1510
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2060 base pairs
 TYPE: NUCLEIC ACID
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: genomic DNA
 PCT-US91-01750-2

Query Match	.2.6%	Score 37.2;	DB 5;	Length 2060;
Best Local Similarity	46.2%;	Pred. No. 0.73;		
Matches 123;	Conservative 0;	Mismatches 143;	Indels 0;	Gaps 0;
Qy	365	GATTATTTGGCATATTCCTCCATCTAAGGATTGAACAGTGTAGTGCGTTATATAAGTTTT	424	
Db	300	GATTAAATAGTGGGAGATTCATTTTTTAAGGATTTTGAAGAATTTCAAGATTCATATGTGTTATT	241	
Qy	425	TGTGCACCAACTCAGAAAGTCGTACATCTTTGAAGTTTGAATTTTCTACTTGCACATTTTAAGT	484	
Db	240	GCTATATATTATTCCTAAAATTTTTAAAAAGTCCTGTGTTATTTGGTGTGGATGTTTATAAT	191	
Qy	485	CCACTTAAATTTGTTTGTGAAGTGAATGTCTACTTTTCAGACACATCTTTTTCCTGCTTCT	544	
Db	180	TCCATATAATCTTTTGTGTTATTCGGTACACAATTTTCTTAATCAANTAAATTTTTTTTTTATT	121	
Qy	545	CTGAGACTCTGCTTTAGTTTGAATCTTTTTTTTGGTCTGTTTTCCTCAGATGCCAATCCCT	604	
Db	120	ATTTAAAAAGCTCAGTATTTTTCGTGTTTTTTTTTTTAAATCTCTGTTACACAAAATAATTGT	61	
Qy	605	ACAATAAGGCTATGCGTTTGCAAAT	630	
Db	60	TACACAAAAAATGTCACCTTTATTAAT	35	

Search completed: June 6, 2003, 16:28:06

Search completed: June
Job time : 253.075 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 6, 2003, 15:14:13 ; Search time 265.151 seconds
(without alignments)
7442.356 Million cell updates/sec

Title: US-09-775-879-20_COPY_143_1552
Perfect score: 1410
Sequence: 1 atggcgccgataatgctta.....tgcattccatggcaacatctg 1410
Scoring table: IDENTITY_NUC
Gap: 10.0, Gapext 1.0

Searched: 870315 seqs, 699768693 residues

Total number of hits satisfying chosen parameters: 1740770

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications_NA:*

```
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %	Match	Length	ID	Description
1	1410	100.0	1889	10	US-09-775-879-20	Sequence 20, Appl
2	587.4	41.7	2925	10	US-09-775-879-22	Sequence 22, Appl
3	309.4	21.3	846	9	US-09-938-842A-421	Sequence 421, Appl
4	125.4	8.3	285	10	US-09-294-093B-1568	Sequence 1568, Ap
5	106.2	7.5	304	10	US-09-294-093B-220	Sequence 220, Appl
6	73	5.2	277	10	US-09-923-876-5279	Sequence 5279, Ap
7	52.8	3.7	446	10	US-09-960-352-3400	Sequence 3400, Ap
8	52.6	3.7	529	10	US-09-983-965-2109	Sequence 2109, Ap
9	52	3.7	9539	9	US-10-239-676-52	Sequence 52, Appl
10	48.4	3.1	12465	9	US-10-239-676-31	Sequence 31, Appl
11	47.2	3.3	686	9	US-10-198-846-5499	Sequence 44, Appl
12	47.2	3.3	6025	9	US-10-239-676-44	Sequence 44, Appl
13	46.8	3.3	2067	10	US-09-880-107-1750	Sequence 1750, Ap
14	46.8	3.3	2158	10	US-09-925-300-588	Sequence 588, Ap
15	46.8	3.3	10046	9	US-10-073-961-458	Sequence 458, Appl
16	46.8	3.3	10046	10	US-09-764-887-458	Sequence 458, Appl
17	46.2	3.3	8333	9	US-10-239-676-114	Sequence 114, Appl
18	45.8	3.2	9539	9	US-10-239-676-51	Sequence 51, Appl
19	45.2	3.2	425	10	US-09-960-352-4010	Sequence 4010, Ap

```

20 44.6 3.2 5690 9 US-10-239-676-73 Sequence 73, Appl
21 44.6 3.2 6544 9 US-10-239-676-46 Sequence 46, Appl
22 44.4 3.1 1267 12 US-10-001-843-45 Sequence 45, Appl
23 44 3.1 516 10 US-09-960-352-5785 Sequence 5785, Ap
24 44 3.1 7657 9 US-10-239-676-185 Sequence 185, Appl
25 43.8 3.1 11812 9 US-10-239-676-210 Sequence 210, Appl
26 43.6 3.1 12405 9 US-10-239-676-35 Sequence 35, Appl
27 43.4 3.1 640681 10 US-09-790-988-1 Sequence 1, Appl
28 43.2 3.1 5898 9 US-10-239-676-180 Sequence 180, Appl
29 43 3.0 2000 9 US-09-938-842A-3497 Sequence 3497, Ap
30 42.8 3.0 432 9 US-09-918-995-13651 Sequence 13651, A
31 42.8 3.0 15732 9 US-10-239-676-95 Sequence 95, Appl
32 42.6 3.0 428 10 US-09-960-352-573 Sequence 573, Appl
33 42.6 3.0 6306 9 US-10-239-676-224 Sequence 224, Appl
34 42.6 3.0 11047 9 US-10-239-676-188 Sequence 188, Appl
35 42.4 3.0 7195 9 US-10-239-676-30 Sequence 30, Appl
36 42.4 3.0 8333 9 US-10-239-676-113 Sequence 113, Appl
37 42.2 3.0 1682 10 US-09-925-302-298 Sequence 298, Appl
38 42.2 3.0 6046 9 US-10-239-676-15 Sequence 15, Appl
39 42 3.0 393 10 US-09-960-352-4582 Sequence 4582, Ap
40 42 3.0 11735 9 US-10-239-676-34 Sequence 34, Appl
41 41.6 3.0 525 9 US-10-198-846-1483 Sequence 1483, Ap
42 41.4 2.9 639 10 US-09-878-574-4316 Sequence 4316, Ap
43 41.4 2.9 5127 9 US-10-239-676-132 Sequence 132, Appl
44 40.8 2.9 994 9 US-10-198-846-6968 Sequence 6968, Ap
45 40.6 2.9 5337 9 US-10-239-676-137 Sequence 137, Appl

```

ALIGNMENTS

```

RESULT 1
US-09-775-879-20
; Sequence 20, Application US/09775879
; Patent No. US20020088822A1
; GENERAL INFORMATION:
; APPLICANT: Choe, Sunghwa
; APPLICANT: Feldmann, Kenneth A.
; TITLE OF INVENTION: Dwf7 MUTANTS
; FILE REFERENCE: 2225-0003
; CURRENT APPLICATION NUMBER: US/09/775,879
; CURRENT FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 60/1179,901
; PRIOR FILING DATE: 2000-02-02
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20
; TYPE: DNA
; ORGANISM: Genomic dwf7 (Arabidopsis)
US-09-775-879-20

```

```

Query Match 100.0%; Score 1410; DB 10; Length 1889;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1410; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGCGCGGATAATGCTTATCTGATGCGAGTTGTTGACGAAACCTCTTTTACACACCA 60
143 ATGGCGCGGATAATGCTTATCTGATGCGAGTTGTTGACGAAACCTCTTTTACACACCA 202
QY 61 ATCGTTCGTGAGTCATCTTTTTCGCGGGAATCTATGGAACCCCTTACCTCATTTTCCTCCAG 120
203 ATCGTTCGTGAGTCATCTTTTTCGCGGGAATCTATGGAACCCCTTACCTCATTTTCCTCCAG 262
QY 121 ACATGGTCCGGAATACCTCCGCGGAACCTTACTATATCTTCTCCGGTTCTCTCTGG 180
263 ACATGGTCCGGAATACCTCCGCGGAACCTTACTATATCTTCTCCGGTTCTCTCTGG 322
QY 181 TCGTCTTACATCTATTACCTTAAATCAACGTTTACCTTCCCAAGGCTCTCGACTTTTCAC 240
323 TCGTCTTACATCTATTACCTTAAATCAACGTTTACCTTCCCAAGGCTCTCGACTTTTCAC 382
QY 241 TTTTGTATTTCACATATTGCTTAATCGCTTTTCTATGTTATCGATTTTTCATATTTAAGGAAGA 300

```



```
QY 735 TCGGCTGGATTCGTATTTTGTATACAGCGCATCTATCTCTTTTCGTGAGTTGGTA 794
Db 61 TTGGTTTTCTATGACCTCTGTATATATGCTATGCTATCTATCTTTGTGGAGTTGGAA 120
QY 795 TTTATTTGGATGACAGAGCTTCATGACATTAAGCCCTCTCTAAGTATCTCCATGCCA 854
Db 121 TTTACTGGATGACAGAGAGTTGCGATGACATAAAACCATATACAAATATCTGCGACAA 180
QY 855 CCCATCATATCTACACAGCAGCAATACACTCTCTCCATTTTCCCGGTAAGTCTTT 909
Db 181 CCCACATATTTACACAGAGGAATAC-TTGTCTCCATTTCTGCTGGACTGCATTT 234

RESULT 6
US-09-923-876-5279
; Sequence 5279, Application US/09923876
; Patent No. US20020013958A1
; GENERAL INFORMATION:
; APPLICANT: Lalguodi, Raghunath V.
; APPLICANT: Kamigaki, Laura Y. (Ito)
; APPLICANT: Shetman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN SEEDLING
; FILE REFERENCE: PL-0012-1 CON
; CURRENT APPLICATION NUMBER: US/09/923,876
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: 09/298,329
; PRIOR FILING DATE: 1999-04-21
; PRIOR APPLICATION NUMBER: 60/085,331
; PRIOR FILING DATE: 1998-05-05
; NUMBER OF SEQ ID NOS: 6332
; SOFTWARE: PERL Program
; SEQ ID NO 5279
; LENGTH: 277
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20020013958A1 700456730H1
US-09-923-876-5279
```

```
Query Match 5.2%; Score 73; DB 10; Length 277;
Best Local Similarity 65.1%; Pred. No. 4e-09;
Matches 123; Conservative 0; Mismatches 65; Indels 1; Gaps 1;

QY 592 AGATGCAATTCCTACAATAAGGCTATGCGTTTGCAAAATGTTTGCGCAATGAAGGCTAT 651
Db 84 AGATGCCATCCCAACAAATGAAGCTATGAAGAAGCAAAATAGCTGTAGCATCCCAAGGCTAT 143
QY 652 GCGATGGTACACTTCTTCCAACTGTCTCCGAGAGTATGATGAACGTGGTTGGACCAA 711
Db 144 GCGTTTTACTGTCTCTTCCAACTTTATCTGAGTATATGATCGAGAGCGGATGGACCG 203
QY 712 ATGTTTGTGACATGACAGAAATTCGGCTGGATTCTGTATTTTGTATCATCGCCATCTA 771
Db 204 GTGTTACTTTATATACAGGAAATGGTTTTTCTCGGTACCTTG-TTATATGGCTATGTA 262
QY 772 TCTTGTTTT 780
Db 263 TCTCATTTT 271
```

```
RESULT 7
US-09-960-352-3400
; Sequence 3400, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 16511.006/37-21(10298)C
```

```
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 3400
; LENGTH: 446
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 15-LIB3058-052-Q1-K1-D11
US-09-960-352-3400

Query Match 3.7%; Score 52.8; DB 10; Length 446;
Best Local Similarity 46.9%; Pred. No. 0.00092;
Matches 165; Conservative 0; Mismatches 187; Indels 0; Gaps 0;

QY 235 TTTCACTTTTGTATTCACATATTCGCTTAATCGCTTTCTATGTTATCGATTTTCAATTTAA 294
Db 95 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 154
QY 295 GGAAGAGGGTTTCTTCGTACTGTACAGATAATTTGGATTTTGGATTTGATGTGGATATGTTCATGTT 354
Db 155 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 214
QY 355 TGCATTTATGATTTTGTGCAATATCTCCATCTAAGGATTAACAGTTAGTGCCTTA 414
Db 215 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 274
QY 415 TATAAGTTTCTGCAACCAATGAGAAAGTCGTACATCTTTGAAGTTGAATTTTCTACTTG 474
Db 275 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 334
QY 475 CCATTTAAGTCACATTAATTTGTTGAAAGTATGTTCTACTTTTCAGACACATCTTTT 534
Db 335 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 394
QY 535 TTTCTGCTCTCTGAGACCTGCTTAGTTTGAATCTTTTGGTCTGTTTT 586
Db 395 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 446
```

```
RESULT 8
US-09-983-965-2109
; Sequence 2109, Application US/09983965
; Patent No. US20020137160A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 37-21(10297)C
; CURRENT APPLICATION NUMBER: US/09/983,965
; CURRENT FILING DATE: 2001-10-26
; PRIOR APPLICATION NUMBER: US 09/465,231
; PRIOR FILING DATE: 1999-12-15
; PRIOR APPLICATION NUMBER: US 60/113,678
; PRIOR FILING DATE: 1998-12-17
; NUMBER OF SEQ ID NOS: 5912
; SEQ ID NO 2109
; LENGTH: 529
; TYPE: DNA
; ORGANISM: Bos taurus
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (335)
; OTHER INFORMATION:
; OTHER INFORMATION: Clone ID: 34-LIB3057-015-Q1-K1-A6
US-09-983-965-2109
```

```
Query Match 3.7%; Score 52.6; DB 10; Length 529;
Best Local Similarity 46.6%; Pred. No. 0.0011;
Matches 166; Conservative 0; Mismatches 190; Indels 0; Gaps 0;
```

```
QY 235 TTTCACATTTGTGATTCACATATGCTTAATCGCTTCTATGTTATCGATTTTCAATTTAA 294
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 154 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 213
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 295 GGAAGA:GGTTCTTCGTTACTGTACAGTAATTTGGATTTGATGATGATGATGATGAT 354
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 214 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 273
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 355 TGCATTTATGATTTATGTTGTCATATCTCCATCTAAGGGATTTGAACAGTTAGTGGCTTA 414
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 274 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 333
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 415 TATAAGTTTGTGCAACCAATGAGAGTCGTCACATCTTTGAAGTTGAATTTTCTACTTG 474
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 334 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 393
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 475 CCATTTAGTCCACATTAATCTGTTGTTGAAGTGAATGCTACTTTCAGACACATCTTT 534
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 394 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 453
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 535 TTCTGCTTCTCTGAGACTCTGCTTGTAGTTTGAATCTTTTGGTCTGTTTGTCTT 590
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 454 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 509
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 9
US-10-239-676-52
; Sequence 52, Application US/10239676
; Publication No. US20030082609A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with Gene Regulation
; FILE REFERENCE: 5013.1003
; CURRENT APPLICATION NUMBER: US/10/239,676
; CURRENT FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: PCT/EP01/03968
; DE 10019058.8
; DE 10019173.8
; DE 10032529.7
; DE 10043826.1
; PRIOR FILING DATE: 2001-04-06
; 2000-04-06
; 2000-04-07
; 2000-06-30
; 2000-09-01
; NUMBER OF SEQ ID NOS: 228
; SEQ ID NO 52
; LENGTH: 9539
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-239-676-52

Query Match 3.7%; Score 52; DB 9; Length 9539;
Best Local Similarity 46.6%; Pred. No. 0.0082;
Matches 166; Conservative 0; Mismatches 190; Indels 0; Gaps 0;

QY 235 TTTCACATTTGTGATTCACATATGCTTAATCGCTTCTATGTTATCGATTTTCAATTTAA 294
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 55 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 114
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 295 GGAAGA:GGTTCTTCGTTACTGTACAGTAATTTGGATTTGATGATGATGATGATGAT 354
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 115 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 174
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 355 TGCATTTATGATTTATGTTGTCATATCTCCATCTAAGGGATTTGAACAGTTAGTGGCTTA 414
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 175 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 234
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 415 TATAAGTTTGTGCAACCAATGAGAGTCGTCACATCTTTGAAGTTGAATTTTCTACTTG 474
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
Db 235 TGTGTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 294
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 475 CCATTTAGTCCACATTAATCTGTTGTTGAAGTGAATGCTACTTTCAGACACATCTTT 534
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 295 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 354
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 535 TTCTGCTTCTCTGAGACTCTGCTTGTAGTTTGAATCTTTTGGTCTGTTTGTCTT 590
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 355 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 410
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 10
US-10-239-676-31
; Sequence 31, Application US/10239676
; Publication No. US20030082609A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with Gene Regulation
; FILE REFERENCE: 5013.1003
; CURRENT APPLICATION NUMBER: US/10/239,676
; CURRENT FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: PCT/EP01/03968
; DE 10019058.8
; DE 10019173.8
; DE 10032529.7
; DE 10043826.1
; PRIOR FILING DATE: 2001-04-06
; 2000-04-06
; 2000-04-07
; 2000-06-30
; 2000-09-01
; NUMBER OF SEQ ID NOS: 228
; SEQ ID NO 31
; LENGTH: 12465
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-239-676-31

Query Match 3.4%; Score 48.4; DB 9; Length 12465;
Best Local Similarity 46.0%; Pred. No. 0.082;
Matches 240; Conservative 0; Mismatches 276; Indels 6; Gaps 2;

QY 103 TTACCTCATTTTCTCCAGACATGGCTCCGAAATTTACCTCGCGGAAACCTACTATCTTC 162
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1661 TTATGGTATTTTTTTAGCGCTAAAGTTTTTGAAGTATTTTTTAAAGTAATGGTTTTTAGTG 1720
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 163 ATCTCCGGTTTCTCTGGTGTCTACATCTATACCTTAAATCAACCGTTTACCTTCCC 222
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1721 TTATTAAGGTTTAAATTTTTTTTGTAGATTTATAGGTTTTTTTATGATTTGATTTGTT 1780
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 223 AAAGTCTCGACTTTTCACTTTTGTATTCATCTATGCTTAATCGCTTTCTATCTTATCGAT 282
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1781 TATTTTTTAACTTTTAAATATATTTATATATATTTTTTTTTTATATATATTTATTTAGTT 1840
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 283 TTTTCAATTTAAGGAAGAGGGTTTCTCGTGTACTGTACAGTAATTTGGATTTGATGCG 342
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1841 TTTTCTATTTTAAAGTGTATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTAT 1896
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 343 ATAGTTCATGTTTGCATTTATTTGATTTATTTGTGATATTTCTCATCAAGGATTTGAACA 402
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1897 AGGATATTTTAAATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTAT 1956
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 403 GTTAGTGGCTTATAAGTTTTTTTGTGCAACCAATGAGAGTCTGTACATCTTTGAAGTTGA 462
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1957 TTAAGTTATTTAGTTAGTTTTTTTGAATATTTATATTTCTTTTATGTTGGTATTTTACT 2016
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 463 ATTTCTACTTCCCATTTAAGTCCACTTAAATTTGTTGTTGAAGTATTTGCTACTTTCA 522
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
Db 2017 TTTTATGATTTTATAGTTAAAT--ATAGAAATTAGAGTAGTTTTTTTTTTGA 2074
QY 523 GACACATCTTTTCTGCTCTGAGACTCTGCTAGTTGAAATCTTTTGTGCTG 582
Db 2075 AATATAGTTTTTTTATTTTTTTTTTTTTTTTTTTTTTTTTTTTATTTATTT 2134
QY 583 TTTTGCTTCAGATGCAATTCCTACAATAAAGGCTATGCGTTT 624
Db 2135 TTTTGCTTAAATGAAAGGTTAGAAAGTTAAATTTGTTGCT 2176

RESULT 11
US-10-198-846-5499
; Sequence 5499, Application US/10198846
; Publication No. US2003009974A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198,846
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5499
; LENGTH: 686
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 2, 4, 22, 33, 44, 149, 196, 204, 205, 206, 210, 213, 220,
; LOCATION: 227, 228, 229, 230, 231, 244, 247, 248, 249, 250, 251, 252,
; LOCATION: 253, 254, 255, 256, 261, 271, 272, 276, 277, 279, 282, 283,
; LOCATION: 284, 285, 286, 291, 292, 294, 295, 297, 298, 299, 302
; OTHER INFORMATION: n = A,T,C or G
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 303, 304, 305, 306, 307, 308, 309, 310, 311, 312, 313, 314,
; LOCATION: 315, 317, 318, 319, 320, 321, 328, 331, 334, 336, 343, 344,
; LOCATION: 347, 348, 349, 350, 351, 353, 355, 359, 360, 362, 363, 368,
; LOCATION: 369, 370, 371, 372, 373, 379, 384, 385, 387, 391, 392
; OTHER INFORMATION: n = A,T,C or G
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 393, 396, 397, 398, 399, 402, 405, 406, 407, 408, 413, 415,
; LOCATION: 416, 426, 430, 433, 434, 438, 440, 441, 442, 445, 448, 463,
; LOCATION: 468, 469, 471, 486, 487, 488, 489, 490, 491, 493, 494, 496,
; LOCATION: 497, 499, 500, 501, 502, 505, 510, 511, 515, 517, 518
; OTHER INFORMATION: n = A,T,C or G
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 520, 527, 528, 529, 535, 536, 539, 557, 559, 562, 565, 567,
; LOCATION: 569, 570, 573, 581, 582, 584, 585, 590, 599, 602, 604, 608,
; LOCATION: 616, 617, 619, 623, 626, 627, 630, 632, 633, 634, 635, 636,
; LOCATION: 637, 638, 641, 643, 644, 648, 651, 653, 655, 658, 661
; OTHER INFORMATION: n = A,T,C or G
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 662, 665, 669, 676, 677
; OTHER INFORMATION: n = A,T,C or G
US-10-198-846-5499
Query Match 3.3%; Score 47.2; DB 9; Length 686;
Best Local Similarity 31.4%; Pred. No. 0.033;
Matches 109; Conservative 0; Mismatches 238; Indels 0; Gaps 0;
QY 243 TTGATTCATATTGCTTAATCGCTTCTATGTTATCGATTTTCAATTTAAGGAGG 302
```

```
Db 257 TTTTNTCTCTTTTNTTNNCTTNNNNCTTNNNNCTTNNNNCTNNNNNNNNNNNT 316
QY 303 GTTCTCTCTGCTACGTACAGTAATTTGGATTGATGCTGATAGTTCATGTTGCATTAA 362
Db 317 NNNNNNTTTTNTNTTNCNTCTTTTNTTNNNNNTNTTNTTNNNNCTTNNNNNNNTT 376
QY 363 TTGATTATTGTGCAATATCTCCATCTAAGGATTAACAGATAGTGGCTTATATAAGTT 422
Db 377 TTTTNTTNTTNTTNNNNNTTNNNNNTTNNNNNTTNTTNTTCTTNCCTTNNNT 436
QY 423 TTTGTGCAACCAATGAGAGTGTGATCATCTTTGAAAGTTGAAATTTCTACTTGCATTAA 482
Db 437 TTTNNNNCTNTTNTTTTTTTTTTTTTTTTCTTNNNNCTTTTTTTTTTTTTNNNNNNNTN 496
QY 483 GTCCACTTAAATTTGTTTGAAGTGTCTACTTTTCAGACACATCTTTTCTGCTT 542
Db 497 NTNNNNNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTT 556
QY 543 CTCTGAGACTCTGCTCTAGTTTGAATCTTTTTTGGTCTGTTTGTCT 589
Db 557 NTNTTNTTNTNNCTNTTTTCTCTNNNTTNTTNTTNTTNTTNTTNTTNTTNTTNT 603

RESULT 12
US-10-239-676-44
; Sequence 44, Application US/10239676
; Publication No. US20030082609A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with Gene Regulation
; FILE REFERENCE: 5013.1003
; CURRENT APPLICATION NUMBER: US/10/239,676
; PRIOR FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: PCT/EP01/03968
; DE 10019058.8
; DE 10019173.8
; DE 10032529.7
; DE 10043826.1
; PRIOR FILING DATE: 2001-04-06
; 2000-04-06
; 2000-04-07
; 2000-06-30
; 2000-09-01
; NUMBER OF SEQ ID NOS: 228
; SEQ ID NO 44
; LENGTH: 6025
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Chemically treated genomic DNA (Homo sapiens)
US-10-239-676-44
Query Match 3.3%; Score 47.2; DB 9; Length 6025;
Best Local Similarity 45.8%; Pred. No. 0.11;
Matches 163; Conservative 0; Mismatches 193; Indels 0; Gaps 0;
QY 235 TTTTCACTTTTGTATTCACACTATGCTTAATCGCTTTCTATGTTATCGATTTTCAATTAA 294
Db 2987 TTTTGTATTGTTATTTTCTTTTGGCGATATAGTAGATTTTAAAAA 3046
QY 295 GGAAGAGGGTTTCTTCTGCTACTGTACATAATTTGGATTGATGGATAGTTCATGTT 354
Db 3047 AAAATGATTAATAACGGATAAGTGGATTTTATATAGAGATTATTTATGTTATTAG 3106
QY 355 TGCATTATTGATTATTTTGTGCATATCTCCATCTAAAGGATTAACAGTATTAGTGCCTTA 414
Db 3107 ATTATTTAAATAATTTTCTGATCGGTTGTTATTTTGTGTTAGAGTGGGTGA 3166
QY 415 TATAAGTTTTTGTGCAACCAATGAGAGTCGTACATCTTTTGAAGTTGAATTTTCTACTG 474
```

Db 3167 GGTTATTTTTCGAGGAGAAAAGATAGATAGATGTTTTATTTTTTTTTTTT 3226
Qy 475 CCATTTLAGTCCACTTAAATTTGTTTGAAGTCATGCTACCTTTCAGACACATCTTT 534
Db 3227 TTTTTLTATT 3286
Qy 535 TTTGCTCTCTCTGAGACTCTGCTTAGTTTGAATCTTTTTTGTGCTGTTTGTCTT 590
Db 3287 TTTTTLTATT 3342

RESULT 13

US-09-880-107-1750
; Sequence 1750, Application US/09880107
; Patent No. US2002042981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scher, Uwe
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 4-921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 1750
; LENGTH: 2067
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 D85181
; US-09-880-107-1750

Query Match 3.3%; Score 46.8; DB 10; Length 2067;
Best Local Similarity 57.5%; Pred. No. 0.078;
Matches 84; Conservative 0; Mismatches 62; Indels 0; Gaps 0;
Qy 1247 TTTTACAGGGCTTGCAATTTCCACCGAGTACCTTTCAGGCTGTACCGCATGTGATA 1306
Db 747 TTTGCAAGTCATGCTTTTCCACCTATTGATGGCTTTCTTCAGAGCTTACCTTACCATA 806
Qy 1307 GCGCTGTTTATAGTCCCAATTCATTTCAACACTCATATAGTCTTTTCTTACATGGAAGCG 1366
Db 807 TACCTTTTATCTTTCCATTACACAAGGTGGTTTATTAAGTCTGTACATCTTGGTTAAT 866
Qy 1367 ATATGGACGGGCAACATCCATGACTG 1392
Db 867 ATCTGGACAATTTCCATTTCATGACGG 892

RESULT 14

US-09-925-300-588
; Sequence 588, Application US/09925300
; Patent No. US20020151681A1
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruben
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: P101
; CURRENT APPLICATION NUMBER: US/09/925,300
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: PatentIn ver. 2.0
; SEQ ID NO 588

; LENGTH: 2158
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-300-588
Query Match 3.3%; Score 46.8; DB 10; Length 2158;
Best Local Similarity 57.5%; Pred. No. 0.08;
Matches 84; Conservative 0; Mismatches 62; Indels 0; Gaps 0;
Qy 1247 TTTTACAGGGCTTGCAATTTCCACCGAGTACCTTTCAGGCTGTACCGCATGTGATA 1306
Db 596 TTTGCAAGTCATGCTTTTCCACCTATTGATGGCTTTCTTCAGAGCTTACCTTACCATA 655
Qy 1307 GCGCTGTTTATAGTCCCAATTCATTTCAACACTCATATAGTCTTTTCTTACATGGAAGCG 1366
Db 656 TACCTTTTATCTTTCCATTACACAAGGTGGTTTATTAAGTCTGTACATCTTGGTTAAT 715
Qy 1367 ATATGGACGGGCAACATCCATGACTG 1392
Db 716 ATCTGGACAATTTCCATTTCATGACGG 741

RESULT 15

US-10-073-961-458
; Sequence 458, Application US/10073961
; Publication No. US20030077602A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P113C1
; CURRENT APPLICATION NUMBER: US/10/073,961
; CURRENT FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: 09/764,887
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/214,886
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/217,487
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,758
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,963
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/217,496
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,447
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/218,290
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/225,757
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/226,868
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 60/216,647
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/225,267
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/216,880
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/225,270
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/251,869
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/235,834
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/234,274
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: 60/234,223
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: 60/228,924

;
; PRIOR FILING DATE: 2000-08-30
; PRIOR APPLICATION NUMBER: 60/224,518
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/236,369
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/224,519
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,964
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/241,809
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/249,299
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/236,327
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/241,785
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/244,617
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 60/225,268
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/236,368
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/251,856
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/251,868
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/229,344
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/234,997
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: 60/229,343
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/229,345
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/229,287
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/229,513
; PRIOR FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: 60/231,413
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/229,509
; PRIOR FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: 60/236,367
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/237,039
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/237,038
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/236,370
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/236,802
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/237,037
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/237,040
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/240,960
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/239,935
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/239,937
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/241,787
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/246,474
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/246,532
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/249,216
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,210
; PRIOR FILING DATE: 2000-11-17

;
; PRIOR APPLICATION NUMBER: 60/226,681
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 60/225,759
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/225,213
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/227,182
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 60/225,214
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/235,836
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/230,438
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/215,135
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: 60/225,266
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/249,218
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,208
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,213
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,212
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,207
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,245
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,244
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,217
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,211
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,215
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,264
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,214
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,297
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/232,400
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/231,242
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/232,081
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/232,080
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/231,414
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/231,244
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/233,064
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/233,063
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/232,397
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/232,399
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/232,401
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/241,808
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,826
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,786
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,221

; PRIOR FILING DATE: 2000-10-20
 ; PRIOR APPLICATION NUMBER: 60/246,475
 ; PRIOR FILING DATE: 2000-11-08
 ; PRIOR APPLICATION NUMBER: 60/231,243
 ; PRIOR FILING DATE: 2000-09-08

Query Match	3.3%	Score 46.8;	DB 9;	Length 10046;
Best Local Similarity	57.5%	Pred. No. 0.19;		
Matches 84;	Conservative 0;	Mismatches 62;	Indels 0;	Gaps 0;

QY	1247	TTTACACGGCTTGCATTTACCCAGTAGACACGGGATACCTTCAGGCTGTACCGCATGTGATA	1306
Db	3736	TTTGCAJGTGATGCTTTTCACCCCTATTTCATGGCTTTCTTCAGAGTCTACCTTACCATATA	3795
QY	1307	GCGCTGTATTATAGTGCCAAATTCATTTCACAACCTCATATAGTCTTTTGTTCATGGAAGCG	1366
Db	3796	TACCCCTTTTATCTTTCCATTACACAAGGTGGTTTATTAACTCTGTACATCTTGGTTAAT	3855
QY	1367	ATATGGJCGGCGAACATCCATGACTG	1392
Db	3856	ATCTGGJCAATTTCCATTCATGACGG	3881

Search completed: June 6, 2003, 18:14:19
 Job time : 271.151 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 6, 2003, 12:27:48 ; Search time 3665.11 Seconds
(without alignments)
6230.541 Million cell updates/sec

Title: US-09-775-879-20_COPY_143_1552

Perfect score: 1410

Sequence: 1 atgacgagcgaataatgctta.....tgcatccatggcaacatctg 1410

Scoring table: IDENTITY_NUC
Gap: 10.0, Gapext 1.0

Searched: 16151066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: an_estba:*

2: an_esthum:*

3: an_estin:*

4: an_estmu:*

5: an_estov:*

6: an_estpl:*

7: an_estro:*

8: an_hic:*

9: gb_est1:*

10: gb_est2:*

11: gb_hic:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_estom:*

17: gb_gss:*

18: em_gss_hum:*

19: em_gss_inv:*

20: em_gss_pln:*

21: em_gss_vrt:*

22: em_gss_fun:*

23: em_gss_man:*

24: em_gss_mus:*

25: em_gss_other:*

26: em_gss_pro:*

27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
c					
1	478	33.9	498	17	AQ011450 F27K17TFC
2	183.6	13.0	583	12	BF425767 sr42f04.y
3	182	12.9	716	10	AW775288 EST334353
4	182	12.9	804	13	BI308419 EST529829
5	181.8	12.9	535	10	AW596303 sj01f11.y
6	181.4	12.9	592	10	AW219366 EST301848

7	173.8	12.3	667	10	BE357415
8	173.4	12.3	621	12	BE123510
9	173.4	12.3	753	13	BI930355
10	170.2	12.1	722	12	EG598293
11	169.6	12.0	574	14	Q803827
12	168.8	12.0	708	10	BE602261
13	168.2	11.9	612	14	Q472907
14	168.2	11.9	769	14	Q802075
15	168	11.9	597	12	EG098248
16	167.2	11.9	708	10	BE601702
17	166.8	11.8	608	14	EQ118369
18	166.6	11.8	782	12	BF624091
19	166.6	11.8	782	12	BF624091
20	166.6	11.8	782	12	BF624091
21	166	11.8	672	14	Q806050
22	165	11.7	668	10	BE404301
23	165	11.7	781	12	BF617574
24	163.4	11.6	658	14	Q246840
25	161.8	11.5	605	13	BI417093
26	161.8	11.5	605	13	BI417093
27	158.2	11.2	517	14	Q703739
28	158	11.2	570	13	EG933493
29	157.2	11.1	652	14	Q766124
30	155	11.0	587	12	BF586850
31	152.8	10.8	586	14	BM903391
32	150.6	10.7	984	11	AY110327
33	146.8	10.4	667	17	BH424776
34	145.8	10.3	661	10	BE413481
35	145.2	10.3	562	10	BE125051
36	145	10.3	567	10	AW429266
37	141	10.0	691	13	BI958272
38	135.4	9.6	840	12	EG320935
39	134.4	9.5	486	13	BI643821
40	132.6	9.4	583	14	Q695873
41	124.8	8.9	541	12	EG322635
42	124.6	8.8	572	10	BE471514
43	120	8.5	586	14	BU012147
44	118.8	8.4	551	10	BE471892
45	116.6	8.3	464	10	AW040702

ALIGNMENTS

RESULT 1
AQ011450/c
LOCUS F27K17TFC IGF Arabidopsis thaliana genomic clone F27K17, DNA
DEFINITION 498 bp DNA linear GSS 29-MAY-1998
sequence.
ACCESSION AQ011450
VERSION AQ011450.1 GI:3166695
KEYWORDS GSS.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
REFERENCE Rounsley,S.D., Suh,E.J., Wible,C., Golden,K., Shatsman,S., Choi,P.,
AUTHORS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Eukaryota; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 498)
Rounsley,S.D., Suh,E.J., Wible,C., Golden,K., Shatsman,S., Choi,P.,
Yu,K., Akinretaye,B., Shen,K., Goonasekaram,S., Millscher,J.,
Adams,M.D. and Venter,J.C.
A BAC End Sequence Database for Identifying Minimal Overlaps in
Arabidopsis Genomic Sequencing. Update 4
Unpublished (1998)
Contact: Steve Rounsley
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: rounsley@tigr.org
Seq primer: M13-21
Class: BAC ends


```
FEATURES
source
High quality sequence stop: 498.
Location/Qualifiers
1. .498
/organism="Arabidopsis thaliana"
/strain="Columbia"
/db_xref="taxon:3702"
/cloned_lib="F27K17"
/sex="hermaphrodite"
/note="Vector: Bel08AII; Site_1: EcoRI; Site_2: EcoRI;
Produced by Thomas Altmann"
BASE COUNT      200 a    95 c    83 g    120 t
ORIGIN
Query Match      33.9%; Score 478; DB 17; Length 498;
Best Local Similarity 98.8%; Pred. No. 6.5e-99;
Matches 492; Conservative 0; Mismatches 5; Indels 1; Gaps 1;
QY 221 CCAAAGTCTCGACTTCTACTTTTGTATTCACACTATTGCTTAATCGCTTTCTATGTTATCG 280
DB 498 CCAAAGTCTCGACTTCTACTTTTGTATTCACACTATTGCTTAATCGCTTTCTATGTTATCG 439
QY 281 ATTTTTCACATTT-AAGGAAGAGGGTTTCTCGTGACTGTACAGTAATTTGGATTGATG 339
DB 438 ATTTTTCATTTAAGACAGAGGGTTTCTCGTGACTGTACAGTAATTTGGATTGATG 379
QY 340 TGGATAGTTCATGTTTGCATTTATGATTATTGTGCAATATTCCTCATCTAAGGGATTGA 399
DB 378 TGAATAGTTCATGTTTGCATTTATGATTATTGTGCAATATTCCTCATCTAAGGGATTGA 319
QY 400 ACAGTTAGTGGCTATATAAGTTTGTGCAACCAATGAGAGTCTACATCTTGAAGT 459
DB 318 ACAGTTAGTGGCTATATAAGTTTGTGCAACCAATGAGAGTCTACATCTTGAAGT 259
QY 460 TGAATTTTCTACTTGCATTTAAGTCCACTTAATTTGTTTGAAGTATGTTCTACTT 519
DB 258 TGAATTTTCTACTTGCATTTAAGTCCACTTAATTTGTTTGAAGTATGTTCTACTT 199
QY 520 TCAGACACATCTTTTCTCTCTCTGAGACTCTCTCTAGTTTGAATCTTTTGTGTT 579
DB 198 TCAGACACATCTTTTCTCTCTCTGAGACTCTCTCTAGTTTGAATCTTAATTTGTT 139
QY 580 CTGTTTGTCTGAGTCAATCTCTACATAAAGGCTATGCGTTTCCAATGTTTGTGCG 639
DB 138 CTGTTTGTCTGAGTCAATCTCTACATAAAGGCTATGCGTTTCCAATGTTTGTGCG 79
QY 640 AATGAAGGCTATGCCATGTTACACTCTCTTCCAACCTGCTCCGAGAGTATGATCAACG 699
DB 78 AATGAAGGCTATGCCATGTTACACTCTCTTCCAACCTGCTCCGAGAGTATGATCAACG 19
QY 700 TGGTTGGACCAATGTTT 717
DB 18 TGGTTGGACCAATGTTT 1
RESULT 2
BF425767
LOCUS
DEFINITION
sr42f04.y1 Gm-cl051 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
Gm-cl051-728 5', similar to TR:Q92729 Q92729 STEROL-C5(6
)-DESATURASE. ;, mRNA sequence.
ACCESSION
BF425767
VERSION
BF425767.1 GI:11413756
KEYWORDS
EST.
SOURCE
soybean.
ORGANISM
Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.
1 (bases 1 to 583)
REFERENCE
Shoemaker, R., Keim, P., Vodkin, L., Erpelding, J., Coryell, V., Khanna
, A., Bolla, B., Maria, M., Hillier, L., Kucaba, T., Martin, J., Beck, C.,
Wyllie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers
, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk
, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann
, R., Waterston, R. and Wilson, R.
Public Soybean EST Project
Unpublished (1999)
Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available through: ResGen, Invitrogen Corp. 2130
South Memorial Parkway Huntsville, AL 35801 For further information
call: (800)-533-4363 or contact via email: ccu@resgen.com
High quality sequence stop: 393.
Location/Qualifiers
1. .383
/organism="Glycine max"
/db_xref="taxon:3847"
/cloned_lib="Gm-cl051"
/issue_type="floral meristematic mRNA"
/note="Vector: pBluescript II SK+; Site_1: EcoRI; Site_2:
XhoI; The cDNA library was constructed from floral
meristematic mRNA provided by Dr. Halina Knap of Clemson
University. Complementary DNA was synthesized from mRNA
using a primer consisting of a poly(dT) sequence with a
XhoI restriction site. EcoRI adapters were ligated to the
blunt-ended cDNA fragments followed by XhoI digestion. The
cDNA fragments were directionally cloned into the
EcoRI-XhoI restriction site of the pBluescript vector. The
ligated cDNA fragments were transformed into DH10B host
cells (GibcoBRL). This library was constructed in the
laboratory of Dr. Randy Shoemaker."
BASE COUNT      131 a    149 c    120 g    183 t
ORIGIN
Query Match      13.0%; Score 183.6; DB 12; Length 583;
Best Local Similarity 71.1%; Pred. No. 1.4e-31;
Matches 243; Conservative 0; Mismatches 99; Indels 0; Gaps 0;
QY 578 GTCTGTTTGTCTTCAGATGCAATTCCTACAATAAGGCTATGCGTTTGCAAAATGTTGTG 637
DB 203 GTTATGTTCCCAAGATGCTATTCCTCTCGAAGAGCCATGCTCTTGCAAAATATCTGTT 262
QY 638 GCATGAAGGCTATGCCATGCTGCTACACTCTCTTCCCACTGTCTCCGAGAGTATGATGAA 697
DB 263 GCTATGAAAGCCATGCTGCTGCTACACTTTCCTTCCCACTGTTTCGGAGTACCTGGTAGAA 322
QY 698 CGTGGTTGGACCAAAATGTTTGTCTAGCATAGACGAATTCGGCTGGAATCTGTATTTGTT 757
DB 323 ACTGGCTGGACAAAGTCTATCTAGTATTAATAATGTTGGTGGCTTCATACCTTGTG 382
QY 758 TACATCGCATCTATCTGTTTTCGTTTGGTATTTGTTATTTGATGCACAGAGCTT 817
DB 383 TATTTAGCAATTTATCTAAGTATTTAGAGTTGGTATTTATTTGATGCACAGAGACTG 442
QY 818 CATGACATTAAGCCTCTCTATAGTATTCATGCCACCCATCATATCTACACACAGCAG 877
DB 443 CACGACATAAAACCCGCTTTTACAAATATCTTCTGCTACCCATCACCATACACATAAACAG 502
QY 878 AATACACTCTCCCATTTTCGGTAAAGTGTTCAGTTTGT 919
DB 503 AACACTCTCTCCCTTGTGTGCTTGGCATTTTCATCTCT 544
RESULT 3
AW775288
LOCUS
DEFINITION
EST334353 DSIL Medicago truncatula cDNA clone pDSIL-1E1, mRNA
AW775288
716 bp mRNA linear EST 07-SEP-2000
EST334353
```

[illegible]

JOURNAL
COMMENT

Unpublished (2000)
Contact: Cathy Ranning
The Institute for Genomic Research
For clone info: please contact Research Genetics, Libraries
Division tel 1-800-711-6195, email cdna@resgen.com
Seq primer: M13F-R.

FEATURES
source

Location/Qualifiers
1. .722
/organism="Solanum tuberosum"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone_lib="CSTS"
/clone="CSTS20L7"
/tissue_type="sprouting eyes from tubers"
/dev_stage="12-14 weeks post harvest"
/lab_host="SOLR"
/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; Various sizes of sprouting eyes (2mm to 15mm) were
taken from tubers. The tubers were incubated at 26C in the
dark for 2-3 weeks prior to sprouting. The eyes were
frozen in liquid nitrogen immediately upon removal from
tubers."

BASE COUNT 191 a 165 c 148 g 218 t

ORIGIN

Query Match 12.1%; Score 170.2; DB 12; Length 722;

Best Local Similarity 69.3%; Pred. No. 1.7e-28;

Matches 232; Conservative 0; Mismatches 103; Indels 0; Gaps 0;

Qy 578 GTCGTGTTTCAGATCAATTCCTACAATAAGGCTATGCGTTTGCAAAATGTTGTG 637

Db 193 GTCTATATCCCAAGATGCATACATCAAGGAAGAGTCTCTTGCAAAATATCAGTT 252

Qy 638 GCATGAGGCTATGCGATGACATCTCTTCCCACTCTCCGAGAGTATGATGAA 697

Db 253 GCTATGAAGCTATGCGGTGTACTGTGCCCTTCCATCACTTCTGAATACATGATGAA 312

Qy 698 CGTGGTTGGACCAAAATGTTTGTAGCATAGACGAATTCGCGTGGATTCGTATTTGTT 757

Db 313 AACGGATGGACCAAAATGTTTGGAGATGAATGATGTTGGATGGCTACTACATCATC 372

Qy 758 TACATGCCATCATCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 817

Db 373 AATGCGGCTATTTATCTGTAATGAGTGGAGTTGGAATCTACTGATGCAAGTTGTTG 432

Qy 818 CATGACATTAAGCTCTCTATAAGTATCTCCATGCCACCATCATATATTTACAACAGCA 877

Db 433 CATGACATAAAACCTCTGTACAATATCTGATGCTACATCATATTTACAACAGCAA 492

Qy 878 AATACATCTCTCCATTTGCGGTAAGTGTGTTTCA 912

Db 493 AACACACTTTCCCGTTGCTGGATTGCCATCCA 527

RESULT 11
BQ03827

LOCUS

WHE2842_D06_G12S Triticum monococcum vernalized apex cDNA library
Triticum monococcum cDNA clone WHE2842_D06_G12, mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

BQ03827 574 bp mRNA linear EST 30-JUL-2002
Triticum monococcum
EST.
Triticum monococcum.
Triticum monococcum

REFERENCE

AUTHORS

TITLE

JOURNAL

1 (bases 1 to 574)
Anderson O.D., Chao, S., Crossman, C., Dubcovsky, J., Echenique, V.,
Lazo, G.R., Pham, J., Rausch, C.J., Stamova, B., Wilson, C. and Woo, J.
The structure and function of the expressed portion of the wheat
genomes - Vernalized apex cDNA library from Triticum monococcum
Unpublished (2002)

COMMENT

Contact: Olin Anderson
US Department of Agriculture, Agriculture Research Service, Pacific
West Area, Western Regional Research Center
800 Buchanan Street, Albany, CA 94710, USA
Tel: 5105595773
Fax: 5105595818
Email: oanderson@pw.usda.gov
Sequences have been trimmed
quality sequence with phred score less than 20
Seq primer: SK primer.

FEATURES
source

Location/Qualifiers
1. .574
/organism="Triticum monococcum"
/cultivar="G3116"
/db_xref="taxon:4568"
/clone="WHE2842_D06_G12"
/clone_lib="Triticum monococcum vernalized apex cDNA
library"
/tissue_type="Vernalized apex"
/dev_stage="One month old plants"
/lab_host="E. coli XL0LR"
/note="Vector: Lambda Uni-ZAP XR, excised phagemid;
Site_1: EcoRI; Site_2: XhoI; One-month old plants were
subjected to vernalization treatment by placing them in
the cold room at 6 C, under 15hr light/9hr dark condition.
Total RNA was prepared from apex tissue extracted from
plants with no cold treatment; and from plants with 2-week
, 4-week and 6-week cold treatment separately. Equal
amount of total RNA was pooled from all four samples, a
cDNA library was made using pooled polyA RNA and cDNA
clones were in vivo excised at the University of
California, Davis (V. Echenique, B. Stamova, J. Dubcovsky
) . Plasmid DNA preparations and DNA sequencing were
performed in the OD Anderson lab (all other authors)."

BASE COUNT
ORIGIN

Query Match 12.0%; Score 169.6; DB 14; Length 574;

Best Local Similarity 70.6%; Pred. No. 2.2e-28;

Matches 226; Conservative 0; Mismatches 94; Indels 0; Gaps 0;

Qy 592 AGATGCAATTCCTACAATAAGGCTATGCGTTTGCAAAATGTTGCGCAATGAAGGCTAT 651

Db 231 AGATGTCGTCCTACAGTAGAAGCTATGAAGAACAATAATTGTCATCAAAAGGCTAT 290

Qy 652 GCATGTACACTCTCTTCCAACTGCTCCGAGAGTATGATGAACGTGTTGGACCAA 711

Db 291 GCCTTCTACTGTGCTCTTCCATCCGATCTGACACATGATGAGAGTGGATGGACAG 350

Qy 712 ATGTTTGTAGCATAGACGAATTCGGCTGGATTCTGTATTTTGTATACATCGCATCTA 771

Db 351 GTGTTTCTTCATATCAGCGAAGTGGTTGGCCCATGTACATTTGCTATGTCTCTATA 410

Qy 772 TCTGTTTTCGTTGAGTTGTTGTTTATTTGATGCACAGAGAGCTTCATGACATTAAGCC 831

Db 411 TCTCATCTTTGAGTTTGGATTATTCTGGATGCACAGAGAGTTGTCATGACATAAGCC 470

Qy 832 TCTCTAAGTATCTCCATGCCACCATCATATCTACAACAGCAGAGATACACATCTCTCC 891

Db 471 ACTATACAGCATCTACATGCAACCCACCACCATTTACAACAGGAGATACCTATCACC 530

Qy 892 ATTTGCGGTAAGTGTGTTTC 911

Db 531 ATTTGCGTGGACTAGCATTC 550

RESULT 12

BQ02261

LOCUS

DEFINITION

ACCESSION

BQ02261 708 bp mRNA linear EST 22-OCT-2001
HVSMEH0098E01f Hordeum vulgare 5-45 DAP spike EST library
HVCDA0009 (5 to 45 DAP) Hordeum vulgare cDNA clone HVSMEH0098E01f,
mRNA sequence.
BQ02261

```
VERSION BE602261.2 GI:13189996
KEYWORDS EST.
SOURCE Hordeum vulgare.
ORGANISM Hordeum vulgare.

REFERENCE
AUTHORS Wing, R., Close, T.J., Kleinhofs, A., Wise, R., Begum, D., Frisch, D., Yu
        Y., Henry, D., Palmer, M., Rambo, T., Simmons, J., Choi, D.W., Penton
        R.D., Close, S.J., Oates, R. and Main, D.
TITLE Development of a genetically and physically anchored EST resource
        for barley genomics: Morex 5-45 DAP spike cDNA library
JOURNAL Unpublished (2001)
COMMENT On Aug 21, 2000 this sequence version replaced gi:9859822.
        Contact: Wing RA
        Clemson University Genomics Institute
        Clemson University
        100 Jordan Hall, Clemson, SC 29634, USA
        Tel: 864 656 7288
        Fax: 864 656 4293
        Email: rwing@clemson.edu
        Total bp bases = 375
        Seq primer: ATTTAACTCTACTAAAGG
        High quality sequence start: 7
        High quality sequence stop: 698.
        Location/Qualifiers
            1..708
                /organism="Hordeum vulgare"
                /cultivar="Morex"
                /db_xref="taxon:4513"
                /clone="HVSMEH009BE01F"
                /clone_lib="Hordeum vulgare 5-45 DAP"
                /HVCNDA0009 (5 to 45 DAP)
                /tissue.type="5-45 DAP Spike"
                /lab_host="SOLR"
                /note="Vector: lambdaZAP; Site_1: EcoRI; Site_2: XhoI;
                Plants were grown in the greenhouse at the University of
                California, Riverside (Fenton, SJ Close, TJ Close). Whole
                spikes with awns trimmed were collected at 5, 10, 15, 20,
                30 and 45 DAP (Fenton). Total RNA was prepared from each
                pool, equal quantities of all six RNA pools were combined,
                poly(A) RNA was purified from the mixture, one primary
                unamplified cDNA library was made, and 1 million pfu were
                in vivo excised to give pBluescript SK(-) cDNA phagemids
                (Choi) in the TJ Close lab at the University of California,
                Riverside. Phagemids were plated and picked at the Clemson
                University Genomics Institute (CUGI) (Begum, Palmer,
                Frisch, Atkins and Wing). Plasmid DNA preparations, DNA
                sequencing and sequence analysis were performed at CUGI
                (Wing, Yu, Frisch, Henry, Simmons, Oates, Rambo, Main).
                The sequence has been trimmed to remove vector sequence
                and contains a minimum of 100 bases of phred value 20 or
                above. For more details on library preparation and
                sequence analysis see
                http://www.genome.clemson.edu/projects/barley. To order
                this clone see http://www.genome.clemson.edu/orders Also
                see clone TJ, Wing R, Kleinhofs A, Wise R (2001)
                Genetically and physically anchored EST resources for
                barley genomics. Barley Genetics Newsletter 31:29-30.
                (http://wheat.pw.usda.gov/ggpages/dgn/31/cover.html)"
BASE COUNT 144 a 210 c 182 g 169 t 3 others
ORIGIN
Query Match 12.0%; Score 168.8; DB 10; Length 708;
Best Local Similarity 71.5%; Pred. No. 3.5e-28;
Matches 221; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

QY 592 AGATCAATTCCTACAATAAGGCTATGCTTGGCAATTTGTTGGCAATGAAGGCTAT 651
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 301 AGATGCTGCTCCACAGTAGAGCTATGAGAGCAATATTTGTCATCAAGGCTAT 360
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 652 GCCATGGTACACTCTTTCCAACTGCTCCGAGAGTATGATTGAACGTGGTGGACCA 711
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```



```
QY 592 AGATG:AAATCTCTACATAAAGGCTATCGGTTTGCATAATGTTTGGCAATGAAGGCTAT 651
Db 199 AGATG:GTGTCCTACAGTAGAAGCTATGAAGAAGCAAAATATGTTGCATCAAAAGCTAT 258
QY 652 GCCAT:GTACACTCTTCTTCAACTGCTCCGAGAGTATGATTGAACGTGGTTGGACCA 711
Db 259 GCCCT:CTACAGTGTCTCTCTCTGATCTGAGCAGCATGATTGAGAGTGGATGGACAG 318
QY 712 ATGTT:TTGCTAGCATAGCAAGATTCCGGCTGGATTCTGTATTGTTTATACCGCATCTA 771
Db 319 GTGTT:TTTTCATATCAGTGAAGTTGGTTGGCCGATGATTGTTGATATGCGCTTTATA 378
QY 772 TCTTG:TTTCTGTTGAGTTTGGTATTATTTGATGACAGAGAGTTCATGACATTAAGCC 831
Db 379 TCTCA:CTTTTGTGGAGTTTGGAAATTTACTGGATGCACAGAGAGTTGTCATACATAAACC 438
QY 832 TCTCT:TAAGATATCTCCATGCCACCCATCATATCTACAACAGCAGAGATACACTCTCTCC 891
Db 439 ACTGT:TAAGACCTTACATGCAACCCACCATTTTACAACAGGAGATAACCTATCACC 498
QY 892 ATTTG:ICGG 900
Db 499 ATTTG:TGG 507
```

```
RESULT 14
B0802075
LOCUS
DEFINITION WHE2821.G07_M132S Triticum monococcum vernalized apex cDNA library
Triticum monococcum cDNA clone WHE2821.G07_M13, mRNA sequence.
```

```
B0802075
B0802075.1 GI:22017044
```

```
EST
Triticum monococcum.
Triticum monococcum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
1 (bases 1 to 769)
Anderson, O.D., Chao, S., Crossman, C., Dubcovsky, J., Echenique, V.,
Lazo, G.R., Pham, J., Raush, C.J., Stamova, B., Wilson, C. and Woo, J.
The structure and function of the expressed portion of the wheat
genomes - Vernalized apex cDNA library from Triticum monococcum
Unpublished (2002)
Contact: Olin Anderson
US Department of Agriculture, Agriculture Research Service, Pacific
West Area, Western Regional Research Center
800 Buchanan Street, Albany, CA 94710, USA
Tel: 510/5959773
Fax: 510/5959818
Email: andersn@pw.usda.gov
Sequencing have been trimmed to remove vector sequence and low
quality sequence with phred score less than 20
Seq primer: SK primer
```

```
FEATURES
source
1. .769
/organism="Triticum monococcum"
/cultivar="G3116"
/db_xref="taxon:4568"
/clone="WHE2821.G07_M13"
/clone_lib="Triticum monococcum vernalized apex cDNA
library"
/tissue_type="Vernalized apex"
/dev_stage="One month old plants"
/lab_host="E. coli XL04"
/notes="Vector: Lambda Uni-ZAP XR, excised phagemid;
Site_1: EcoRI; Site_2: XhoI; One-month old plants were
subjected to vernalization treatment by placing them in
the cold room at 6 C, under 15hr light/9hr dark condition.
Total RNA was prepared from apex tissue extracted from
plants with no cold treatment; and from plants with 2-week
, 4-week and 6-week cold treatment separately. Equal
amount of total RNA was pooled from all four samples, a
```

```
cdna library was made using pooled polyA RNA and cDNA
clones were in vivo excised at the University of
California, Davis (V. Echenique, B. Stamova, J. Dubcovsky
). Plasmid DNA preparations and DNA sequencing were
performed in the OD Anderson lab (all other authors). "
```

```
BASE COUNT 163 a 225 c 200 g 181 t
ORIGIN
```

```
Query Match 11.9%; Score 168.2; DB 14; Length 769;
Best Local Similarity 71.5%; Pred. No. 4.9e-28;
Matches 221; Conservative 0; Mismatches 88; Indels 0; Gaps 0;
QY 592 AGATGCAATCTCTACATAAAGGCTATCGGTTTGCATAATGTTTGGCAATGAAGGCTAT 651
Db 256 AGATGCTGTCCCTACAGTAGAAGCTATGAAGAAGCAAAATATGTTGCATCAAAAGGCTAT 315
QY 652 GCCATGGTACACTCTTCTTCAACTGCTCCGAGAGTATGATTGAACGTGGTTGGACCA 711
Db 316 GCCTTTCTACTGTGCTCTTCCATCGTATCTGAGCAGCATGATTGAGAGTGGATGGACAG 375
QY 712 ATGTTTCTAGCATAGCAAGATTCCGGCTGGATTCTGTATTGTTTATACCGCATCTA 771
Db 376 GTGTTTCTTTCATATCAGTGAAGTTGGTTGGCCGATGATTGTTGATGTCCTCTATA 435
QY 772 TCTTGTTTCTGTTGAGTTTGGTATTATTTGATGACAGAGAGTTCATGACATTAAGCC 831
Db 436 TCTCATCTTTCTGGAGTTTGGAAATTTACTGGATGCACAGAGAGTTGTCATACATAAAGCC 495
QY 832 TCTCTAAGATATCTCCATGCCACCCATCATATCTACAACAGCAGAGATACACTCTCTCC 891
Db 496 ACTATACAGCATCTACATGCAACCCACCATTTTACAACAGGAGATAACCTATCACC 555
QY 892 ATTTGCGG 900
Db 556 ATTTGCTGG 564
```

```
RESULT 15
B098248
LOCUS
DEFINITION EST462767 sprouting eyes/shoots Solanum tuberosum cDNA clone
cSTC2K9 5' sequence, mRNA sequence.
```

```
B098248
B098248.1 GI:12588283
```

```
EST.
Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
1 (bases 1 to 597)
van der Hoeven, R.S., Bezzerides, J., Cho, J., Utterback, T., Hansen
, C.L., Bougri, O., Buell, C.R., Ronning, C.M., Tanksley, S.D. and Baker
, B.
```

```
Generation of ESTs from potato sprouting eyes/shoots
```

```
Unpublished (2001)
Contact: Cathy Ronning
The Institute for Genomic Research
For clone info: please contact Research Genetics, Libraries
Division tel 1-800-711-6195, email cdna@resgen.com.
Location/Qualifiers
1. .597
```

```
FEATURES
source
/organism="Solanum tuberosum"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="cSTC2K9"
/clone_lib="sprouting eyes/shoots"
/tissue_type="sprouting tubers"
/dev_stage="12-14 weeks post harvest"
/lab_host="SOLR"
/notes="vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; Various sizes of sprouting eyes (2 mm to 15 mm) were
taken from tubers. The tubers were incubated at 26C in
```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 6, 2003, 10:02:20 ; Search time 54 Seconds
(without alignments)

7506.661 Million cell updates/sec

Title: US-09-775-879-20_COPY_143_322

Perfect score: 180

Sequence: 1 atgcggcggaatgtta.....tcattcgggttctctg 180

Scoring table: IDENTITY_NUC
Gapol: 10.0 , Gapext 1.0

Searched: 2185139 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_101002.*

- 1: /SID22/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
- 2: /SID22/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
- 3: /SID22/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
- 4: /SID22/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
- 5: /SID22/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
- 6: /SID22/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
- 7: /SID22/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
- 8: /SID22/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
- 9: /SID22/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
- 10: /SID22/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
- 11: /SID22/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
- 12: /SID22/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
- 13: /SID22/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
- 14: /SID22/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
- 15: /SID22/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
- 16: /SID22/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
- 17: /SID22/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
- 18: /SID22/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
- 19: /SID22/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
- 20: /SID22/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
- 21: /SID22/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
- 22: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
- 23: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
- 24: /SID22/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	180	100.0	1164	21 AAC41897	Arabidopsis thalia
2	54.8	30.4	1381	22 AAH50964	Sterol C5 desaturase
3	32.6	18.1	1830121	17 AAT42063	Haemophilus influenzae
4	31.8	17.7	1283	22 AAK66079	Human immune/haema
5	31.8	17.7	1283	22 AAK66083	Human immune/haema
6	31.8	17.7	1283	22 AAK66084	Human immune/haema
7	30.2	16.8	311	22 AAH36514	Human colon cancer
8	30.2	16.8	2667	21 ABQ69295	Listeria innocua D
9	30.2	16.8	2709	24 ABQ67863	Listeria innocua D

c 10	30.2	16.8	81905	24 ABQ65244	Listeria innocua D
c 11	30.2	16.8	82689	24 AAK67198	Listeria innocua p
c 12	30	16.7	4775	22 AAK76537	Human immune/haema
c 13	30	16.7	4776	22 AAK76538	Human immune/haema
c 14	30	16.7	11316	22 AAS27731	DNA encoding novel
c 15	29.6	16.4	773	22 AAH03702	Human cDNA clone (
c 16	29.6	16.4	919	22 AAL37315	Human musculoskele
c 17	29.6	16.4	2310	22 AAH15301	Human cDNA sequenc
c 18	29.6	16.4	3810	24 ABN97214	Gene #3712 used to
c 19	29.2	16.2	1403	20 AAX86970	Truncated plastid
c 20	29.2	16.2	2381	24 ABN59690	Novel human coding
c 21	29.2	16.2	2703	21 AAN70093	Plasmodium falcipa
c 22	28.8	16.0	2114	14 AAQ39178	Truncated PVX repl
c 23	28.8	16.0	2208	14 AAQ39180	Truncated PVX repl
c 24	28.8	16.0	4455	14 AAQ39177	PVX ORF1 encoding
c 25	28.6	15.9	1884	23 ABL13343	Drosophila melanog
c 26	28.4	15.8	1638	20 AAX16150	NADH:H2O oxidase a
c 27	28.2	15.7	87350	18 AAX83003	Human WRN genomic
c 28	28	15.6	2192	24 AAD22130	Arabidopsis thalia
c 29	28	15.6	3647	23 ABL02676	Drosophila melanog
c 30	28	15.6	35100	20 AAV73804	KSHV LTR DNA (nucl
c 31	28	15.6	37544	21 AAA50029	Cosmid cHRIM5 enco
c 32	28	15.6	137507	19 AAV19941	KSHV long unique c
c 33	28	15.6	1664976	19 AAV21209	Methanococcus jann
c 34	27.8	15.4	1977	22 AAH68280	C glutamicum codin
c 35	27.8	15.4	309400	22 AAH68534	C glutamicum codin
c 36	27.6	15.3	805	21 AAC77362	Human ORFX ORF291?
c 37	27.6	15.3	820	22 AAH05833	Human cDNA clone (
c 38	27.6	15.3	1003	22 AAH08426	Human guanine exch
c 39	27.6	15.3	1003	22 AAK52823	Human polynucleoti
c 40	27.6	15.3	1210	22 AA160003	Human polynucleoti
c 41	27.6	15.3	1897	22 AAK51839	Human polynucleoti
c 42	27.6	15.3	1897	22 AA158217	Human polynucleoti
c 43	27.6	15.3	2276	21 AAAS9498	DNA encoding Smad1
c 44	27.6	15.3	2397	22 AAH14389	Human cDNA sequenc
c 45	27.6	15.3	2720	22 AAS33176	DNA encoding human

ALIGNMENTS

RESULT 1

AAC41897

ID AAC41897 standard; DNA; 1164 BP.

XX AC AAC41897;

XX AC AAC41897;

XX DT 17-OCT-2000 (first entry)

XX DT 17-OCT-2000 (first entry)

DE Arabidopsis thaliana DNA fragment SEQ ID NO: 33540.

DE Hybridisation assay; genetic mapping; gene expression control;

KW protein identification; signal transduction pathway;

KW metabolic pathway; promoter; termination sequence; ss.

XX Arabidopsis thaliana.

OS Arabidopsis thaliana.

PN EP1033405-A2.

XX EP1033405-A2.

XX 06-SEP-2000.

XX 06-SEP-2000.

XX 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-2000; 2000EP-0301439.

PR 05-MAR-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0121825.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.
 PR 21-APR-1999; 99US-0130449.
 PR 23-APR-1999; 99US-0130510.
 PR 23-APR-1999; 99US-0130891.
 PR 28-APR-1999; 99US-0131149.
 PR 30-APR-1999; 99US-0132048.
 PR 30-APR-1999; 99US-0132407.
 PR 04-MAY-1999; 99US-0132484.
 PR 05-MAY-1999; 99US-0132485.
 PR 06-MAY-1999; 99US-0132486.
 PR 06-MAY-1999; 99US-0132487.
 PR 07-MAY-1999; 99US-0132863.
 PR 11-MAY-1999; 99US-0134256.
 PR 14-MAY-1999; 99US-0134218.
 PR 14-MAY-1999; 99US-0134219.
 PR 14-MAY-1999; 99US-0134221.
 PR 14-MAY-1999; 99US-0134370.
 PR 18-MAY-1999; 99US-0134768.
 PR 19-MAY-1999; 99US-0134941.
 PR 20-MAY-1999; 99US-0135124.
 PR 21-MAY-1999; 99US-0135353.
 PR 24-MAY-1999; 99US-0135359.
 PR 25-MAY-1999; 99US-0136021.
 PR 27-MAY-1999; 99US-0136392.
 PR 28-MAY-1999; 99US-0136782.
 PR 01-JUN-1999; 99US-0137222.
 PR 03-JUN-1999; 99US-0137528.
 PR 04-JUN-1999; 99US-0137502.
 PR 07-JUN-1999; 99US-0137724.
 PR 08-JUN-1999; 99US-0138094.
 PR 10-JUN-1999; 99US-0138540.
 PR 10-JUN-1999; 99US-0138847.
 PR 14-JUN-1999; 99US-0139119.
 PR 16-JUN-1999; 99US-0139452.
 PR 16-JUN-1999; 99US-0139453.
 PR 17-JUN-1999; 99US-0139492.
 PR 18-JUN-1999; 99US-0139454.
 PR 18-JUN-1999; 99US-0139455.
 PR 18-JUN-1999; 99US-0139456.
 PR 18-JUN-1999; 99US-0139457.
 PR 18-JUN-1999; 99US-0139458.
 PR 18-JUN-1999; 99US-0139459.
 PR 18-JUN-1999; 99US-0139460.
 PR 18-JUN-1999; 99US-0139461.
 PR 18-JUN-1999; 99US-0139462.
 PR 18-JUN-1999; 99US-0139463.
 PR 18-JUN-1999; 99US-0139750.
 PR 18-JUN-1999; 99US-0139763.
 PR 21-JUN-1999; 99US-0139817.
 PR 22-JUN-1999; 99US-0139899.
 PR 23-JUN-1999; 99US-0140353.
 PR 23-JUN-1999; 99US-0140354.
 PR 24-JUN-1999; 99US-0140695.
 PR 28-JUN-1999; 99US-0140823.
 PR 29-JUN-1999; 99US-0140991.
 PR 30-JUN-1999; 99US-0141287.
 PR 01-JUL-1999; 99US-0141842.
 PR 01-JUL-1999; 99US-0142154.
 PR 02-JUL-1999; 99US-0142055.
 PR 06-JUL-1999; 99US-0142390.
 PR 08-JUL-1999; 99US-0142803.
 PR 09-JUL-1999; 99US-0142920.
 PR 12-JUL-1999; 99US-0142977.
 PR 13-JUL-1999; 99US-0143542.
 PR 14-JUL-1999; 99US-0143624.
 PR 15-JUL-1999; 99US-0144005.
 PR 16-JUL-1999; 99US-0144085.
 PR 16-JUL-1999; 99US-0144086.
 PR 19-JUL-1999; 99US-0144325.
 PR 19-JUL-1999; 99US-0144331.
 PR 19-JUL-1999; 99US-0144332.
 PR 19-JUL-1999; 99US-0144333.
 PR 19-JUL-1999; 99US-0144334.

PR 19-JUL-1999; 99US-0144335.
 PR 20-JUL-1999; 99US-0144352.
 PR 20-JUL-1999; 99US-0144632.
 PR 20-JUL-1999; 99US-0144884.
 PR 21-JUL-1999; 99US-0144814.
 PR 21-JUL-1999; 99US-0145086.
 PR 21-JUL-1999; 99US-0145088.
 PR 22-JUL-1999; 99US-0145085.
 PR 22-JUL-1999; 99US-0145087.
 PR 22-JUL-1999; 99US-0145089.
 PR 22-JUL-1999; 99US-0145192.
 PR 23-JUL-1999; 99US-0145145.
 PR 23-JUL-1999; 99US-0145218.
 PR 23-JUL-1999; 99US-0145224.
 PR 26-JUL-1999; 99US-0145276.
 PR 27-JUL-1999; 99US-0145913.
 PR 27-JUL-1999; 99US-0145918.
 PR 27-JUL-1999; 99US-0145919.
 PR 28-JUL-1999; 99US-0145951.
 PR 02-AUG-1999; 99US-0146386.
 PR 02-AUG-1999; 99US-0146388.
 PR 02-AUG-1999; 99US-0146389.
 PR 03-AUG-1999; 99US-0147038.
 PR 04-AUG-1999; 99US-0147204.
 PR 04-AUG-1999; 99US-0147302.
 PR 05-AUG-1999; 99US-0147192.
 PR 05-AUG-1999; 99US-0147260.
 PR 06-AUG-1999; 99US-0147303.
 PR 06-AUG-1999; 99US-0147416.
 PR 09-AUG-1999; 99US-0147493.
 PR 09-AUG-1999; 99US-0147935.
 PR 10-AUG-1999; 99US-0148171.
 PR 11-AUG-1999; 99US-0148319.
 PR 12-AUG-1999; 99US-0148341.
 PR 13-AUG-1999; 99US-0148565.
 PR 13-AUG-1999; 99US-0148684.
 PR 16-AUG-1999; 99US-0149368.
 PR 17-AUG-1999; 99US-0149175.
 PR 18-AUG-1999; 99US-0149426.
 PR 20-AUG-1999; 99US-0149722.
 PR 20-AUG-1999; 99US-0149723.
 PR 20-AUG-1999; 99US-0149929.
 PR 23-AUG-1999; 99US-0149902.
 PR 23-AUG-1999; 99US-0149930.
 PR 25-AUG-1999; 99US-0150566.
 PR 26-AUG-1999; 99US-0150884.
 PR 27-AUG-1999; 99US-0151065.
 PR 27-AUG-1999; 99US-0151066.
 PR 27-AUG-1999; 99US-0151080.
 PR 30-AUG-1999; 99US-0151303.
 PR 31-AUG-1999; 99US-0151438.
 PR 01-SEP-1999; 99US-0151930.
 PR 07-SEP-1999; 99US-0152363.
 PR 10-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0153758.
 PR 15-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154039.
 PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155486.
 PR 24-SEP-1999; 99US-0156559.
 PR 28-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159295.
 PR 14-OCT-1999; 99US-0159329.

PR 14-OCT-1999; {90S-0159330.
 PR 14-OCT-1999; {90S-0159331.
 PR 14-OCT-1999; {90S-0159637.
 PR 14-OCT-1999; {90S-0159638.
 PR 18-OCT-1999; {90S-0159584.
 PR 21-OCT-1999; {90S-0160741.
 PR 21-OCT-1999; {90S-0160767.
 PR 21-OCT-1999; {90S-0160768.
 PR 21-OCT-1999; {90S-0160770.
 PR 21-OCT-1999; {90S-0160814.
 PR 21-OCT-1999; {90S-0160815.
 PR 22-OCT-1999; {90S-0160980.
 PR 22-OCT-1999; {90S-0160981.
 PR 22-OCT-1999; {90S-0160989.
 PR 25-OCT-1999; {90S-0161404.
 PR 25-OCT-1999; {90S-0161405.
 PR 25-OCT-1999; {90S-0161406.
 PR 26-OCT-1999; {90S-0161359.
 PR 26-OCT-1999; {90S-0161360.
 PR 26-OCT-1999; {90S-0161361.
 PR 28-OCT-1999; {90S-0161920.
 PR 28-OCT-1999; {90S-0161992.
 PR 28-OCT-1999; {90S-0161993.
 PR 29-OCT-1999; {90S-0162142.
 PR 29-OCT-1999; {90S-0162142.
 Query Match 100.0%; Score 180; DB 21; Length 1164;
 Best Local Similarity 100.0%; Pred. No. 3.9e-53;
 Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ATGGCGCGGATAAGCTTAATCTGATGAGTGTGAGCAACCTCTTTTACACCGA 60
 Db 146 ATGGCGCGGATAAGCTTAATCTGATGAGTGTGAGCAACCTCTTTTACACCGA 205
 -QY 61 ATCGTTTGAGTCATCTTTGCGCGGGAATCTATGAGTGTGAGCAACCTCTTTTACACCGA 120
 Db 206 ATCGTTTGAGTCATCTTTGCGCGGGAATCTATGAGTGTGAGCAACCTCTTTTACACCGA 265
 QY 121 ACATGGTCCGAAATACCTCCCGGAACCTTACTATCTTCTCTCTCTCTCTCTG 180
 Db 266 ACATGGTCCGAAATACCTCCCGGAACCTTACTATCTTCTCTCTCTCTCTG 325
 RESULT 2
 AAH50964
 ID AAH50964 standard; cdNA; 1381 BP.
 XX AC
 AC AAH50964;
 DT 28-AUG-2001 (first entry)
 XX DE Sterol C5 desaturase long partial clone nucleotide sequence.
 XX KW Moss; Physcomitrella patens; lipid metabolism related protein; LMRP;
 KW lipid biosynthesis; lipid modification; lipid degradation; cofactor;
 KW fatty acid transport; genetic engineering; fatty acid; enzyme; plant;
 KW microorganism; polyunsaturated fatty acid; oilseed plant; maize; wheat;
 KW biotic stress tolerance; abiotic stress tolerance; rye; oat; triticale;
 KW rice; barley; soybean; peanut; cotton; rapeseed; canola; manihot;
 KW pepper; sunflower; tagetes; potato; tobacco; eggplant; tomato; Vicia;
 KW pea; alfalfa; coffee; cacao; tea; Salix; oil palm; coconut;
 KW perennial grass; forage crop; ss.
 XX OS Physcomitrella patens.
 XX PN WO200138484-A2
 XX PD 31-MAY-2001.
 XX PF 22-NOV-2000; 2000WO-EP11615.
 XX PR 25-NOV-1999; 99WO-EP09108.
 XX PR (BADI) BASF PLANT SCI GMBH.
 PA

XX Lerchl J, Renz A, Ehrhardt T, Reindl A, Cirpus P, Bischoff F;
 PI Frank M, Freund A, Duwenig E, Schmidt R, Reski R;
 XX WPI; 2001-367669/38.
 DR Nucleic acids encoding lipid metabolism related proteins from
 PT Physcomitrella patens useful to produce fine chemicals in modified
 PT organisms, particularly polyunsaturated fatty acids in oilseed plants -
 XX Claim 7; Page 106; 120pp; English.
 XX The present invention describes isolated nucleic acid sequences which
 CC encode lipid metabolism related proteins (LMRP). The LMRP nucleic acids
 CC can be used to modify lipids and fatty acids, cofactors and enzymes in
 CC microorganisms and plants, particularly to produce polyunsaturated fatty
 CC acids, and are especially useful in oilseed plants. The nucleic acids
 CC may also confer biotic or abiotic stress tolerance, particularly to
 CC maize, wheat, rye, oat, triticale, rice, barley, soybean, peanut,
 CC cotton, rapeseed, canola, manihot, pepper, sunflower, tagetes, potato,
 CC tobacco, eggplant, tomato, Vicia species, pea, alfalfa, coffee, cacao,
 CC tea, Salix species, oil palm, coconut, perennial grasses and forage
 CC crops. AAH50878 to AAH50882 represent primers used in the exemplification
 CC of the present invention. AAH50883 to AAH50968 represents LMRP nucleotide
 CC sequences, and AAH50843 to AAH50928 represent LMRP protein sequences,
 CC given in the present invention.
 XX SQ Sequence 1381 BP; 323 A; 331 C; 351 G; 376 T; 0 other;
 Query Match 30.4%; Score 54.8; DB 22; Length 1381;
 Best Local Similarity 59.7%; Pred. No. 5.1e-09;
 Matches 92; Conservative 0; Mismatches 62; Indels 0; Gaps 0;
 QY 27 GCAGTTGTGAGCAACCTCTTTTACACCGAATCTCTGAGTCATCTTTGCGCGG 86
 Db 231 GCTTTCTGAGGAGACCCCGGTGATCAACGATCTGGTCTCGGGCCCTGGCTGCCCTC 290
 QY 87 GAATCTATGGGAACCTTACCTCATTTCTCCAGACATGGCTCCGAAATACCTGCCCGG 146
 Db 291 CTCTGTCGCGACTCCATTCGCCACACATTCGACGACATGGCTGCGGAACTAGCTCGCGG 350
 QY 147 AACCTTACTATCTCATCTCCGGTTCTCTCTGG 180
 Db 351 CATGCTTTGTATTTCTCTCGGTGCGCTGTGG 384
 RESULT 3
 AAT42063
 ID AAT42063 standard; DNA; 1830121 BP.
 XX AC
 AC AAT42063;
 DT 14-SEP-1999 (first entry)
 XX DE Haemophilus influenzae complete genome sequence.
 XX Genome; bacterium; Haemophilus influenzae; computer readable medium;
 KW expression modulating fragment; regulation; gene expression; vector;
 KW organism; open reading frame; ORF; ds.
 XX OS Haemophilus influenzae.
 XX PN WO9633276-A1.
 XX PD 24-OCT-1996.
 XX PF 22-APR-1996; 96WO-US05320.
 XX PR 07-JUN-1995; 95US-0487429.
 XX PR 21-APR-1995; 95US-0426787.
 XX PR 07-JUN-1995; 95US-0476102.
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA

PA (UYJO) UNIV JOHNS HOPKINS.
XX Adams MD, Fleischmann RD, Smith HO, Venter JC, White O;
PI WPI; 1996-485782/48.
XX Haemophilus influenzae Rd genome recorded on computer readable
PT medium - useful for identifying commercially important nucleic acid
PT fragments by homology searching
XX ClalM 1; Page 77.2-77.1091; 1291pp; English.
XX This sequence represents the complete genome sequence of the bacterium
CC Haemophilus influenzae strain Rd. The invention relates to a computer
CC readable medium (CRM) having recorded upon it the complete H.influenzae
CC nucleotide sequence (I), a representative fragment of (I) or a nucleotide
CC sequence at least 99% identical to (I). By providing the full-length
CC genomic sequence in a computer readable form, it is possible to identify
CC commercially important nucleic acid fragments and expression modulating
CC fragments (EMFs) of the Haemophilus genome. The EMFs can be used to
CC regulate the expression of a nucleic acid molecule. Vectors and altered
CC organisms comprising the predicted ORFs can be used to produce any of the
CC polypeptide fragments of the H. influenzae Rd genome.
XX Sequence 1830121 BP; 567399 A; 350615 C; 347389 G; 564036 T; 682 other;
SQ

Query Match 18.1%; Score 32.6; DB 17; Length 1830121;
Best Local Similarity 54.6%; Pred. No. 5.3;
Matches 65; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

QY 59 GAATGCTTCTGAGTCATCTTTGCGCGGCGGAATCTATGGAACCCCTTACCTCATTTTCCTCC 118
DB 1431803 GAGCCATACAGACATCTTTGGTGCAGAAATACCCGACATCATTAATTTATTACCCA 1431862

QY 119 AGACATGGCTCCGAATTAACCTCGCGGGAACCTACTATCTCTCCGCTTCCCTC 177
DB 1431863 AATATTGATGAAGATTAGCTTCGGGTAAACATAAAGCGTTATACCCGTTTCCCG 1431921

RESULT 4
AAK66079/c
ID AAK66079 standard; DNA; 1283 BP.
AC AAK66079;
XX
XX 06-NOV-2001 (first entry)
DT Human immune/haematopoietic antigen genomic sequence SEQ ID NO:20891.
XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis; ds.
XX Homo sapiens.
XX WO200157182-A2.
XX
XX 09-AUG-2001.
XX 17-JAN-2001; 2001WO-US01354.
XX 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0228287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 13-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240560.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.

PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 03-SEP-2000; 2000US-0229509.
PR 03-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 25-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-024617.
PR 01-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.

PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249266.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 01-DEC-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 06-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX (HUMA-) HUMAN GENOME SCI INC.
PA
XX Rosen CA, Barash SC, Ruben SM;
PI WPI; 2001-483426/52.
XX
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and
PT metastasis -
XX
XX Disclosure; SEQ ID NO 20895; 3071pp + Sequence Listing; English.
PS
XX AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
CC amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patients own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting
CC the nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/hematopoietic-related diseases, especially
CC cancers and cancer metastases of hematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/hematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169
CC represent sequences used in the exemplification of the present invention.
XX
XX
SQ Sequence 1283 BP; 418 A; 221 C; 261 G; 383 T; 0 other;

Query Match 17.7%; Score 31.8; DB 22; Length 1283;
Best Local Similarity 54.8%; Pred. No. 0.62;
Matches 63; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

QY 9 GGATATGCTATCTGATGCTTTCTTGTACGAAACCTCTTTTACAAACGAATCGTTCT 68
||||| | | | | | | | | | | | | | | | | | | | | | | | | | | |

Db 783 GGATTAAGATTATTAATGCTGCTGCTGAAGGGAACCACTTTTCATACACAAATGTGATC 724
QY 69 GAGTCATCTTTGCGCGGAATCTATGGGAACCCCTTACCTCATTTTCTCCAGACA 123
Db 723 TTTTTCCTTTTCARAGGCTAATACACAGTACATTTCATATAACTAATCCCAACA 669

RESULT 6
AAK66084/c
ID AAK66084 standa:d; DNA; 1283 BP.
XX AAK66084;
AC
XX
DT 06-NOV-2001 (first entry)
XX
DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:20896.
XX
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis; ds.
XX
OS Homo sapiens.
XX
PN WO200157182-A2.
XX
PD 09-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01354.
XX
XX 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226688.
PR 23-AUG-2000; 2000US-0227182.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 13-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 20-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.

PR 30-JUN-2000; 2000US-0215135.
 PR 07-JUL-2000; 2000US-0216647.
 PR 07-JUL-2000; 2000US-0216880.
 PR 11-JUL-2000; 2000US-0217487.
 PR 11-JUL-2000; 2000US-0217496.
 PR 14-JUL-2000; 2000US-0218290.
 PR 26-JUL-2000; 2000US-0220963.
 PR 26-JUL-2000; 2000US-0220964.
 PR 14-AUG-2000; 2000US-0224518.
 PR 14-AUG-2000; 2000US-0224519.
 PR 14-AUG-2000; 2000US-0225213.
 PR 14-AUG-2000; 2000US-0225214.
 PR 14-AUG-2000; 2000US-0225266.
 PR 14-AUG-2000; 2000US-0225267.
 PR 14-AUG-2000; 2000US-0225268.
 PR 14-AUG-2000; 2000US-0225270.
 PR 14-AUG-2000; 2000US-0225447.
 PR 14-AUG-2000; 2000US-0225757.
 PR 14-AUG-2000; 2000US-0225758.
 PR 14-AUG-2000; 2000US-0225759.
 PR 18-AUG-2000; 2000US-0226279.
 PR 22-AUG-2000; 2000US-0226681.
 PR 22-AUG-2000; 2000US-0226686.
 PR 22-AUG-2000; 2000US-0227182.
 PR 23-AUG-2000; 2000US-0227009.
 PR 30-AUG-2000; 2000US-0228924.
 PR 01-SEP-2000; 2000US-0229287.
 PR 01-SEP-2000; 2000US-0229343.
 PR 01-SEP-2000; 2000US-0229344.
 PR 01-SEP-2000; 2000US-0229345.
 PR 05-SEP-2000; 2000US-0229509.
 PR 05-SEP-2000; 2000US-0229513.
 PR 06-SEP-2000; 2000US-0230437.
 PR 06-SEP-2000; 2000US-0230438.
 PR 08-SEP-2000; 2000US-0231242.
 PR 08-SEP-2000; 2000US-0231243.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231413.
 PR 08-SEP-2000; 2000US-0231414.
 PR 08-SEP-2000; 2000US-0232080.
 PR 08-SEP-2000; 2000US-0232081.
 PR 12-SEP-2000; 2000US-0231988.
 PR 14-SEP-2000; 2000US-0232397.
 PR 14-SEP-2000; 2000US-0232398.
 PR 14-SEP-2000; 2000US-0232399.
 PR 14-SEP-2000; 2000US-0232400.
 PR 14-SEP-2000; 2000US-0232401.
 PR 14-SEP-2000; 2000US-0233063.
 PR 14-SEP-2000; 2000US-0233064.
 PR 14-SEP-2000; 2000US-0233065.
 PR 21-SEP-2000; 2000US-0234223.
 PR 21-SEP-2000; 2000US-0234274.
 PR 25-SEP-2000; 2000US-0234997.
 PR 25-SEP-2000; 2000US-0234998.
 PR 26-SEP-2000; 2000US-0235484.
 PR 27-SEP-2000; 2000US-0235484.
 PR 27-SEP-2000; 2000US-0235836.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.
 PR 29-SEP-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 02-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239935.
 PR 13-OCT-2000; 2000US-0239937.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.

PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 20-OCT-2000; 2000US-0241826.
 PR 01-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 11-DEC-2000; 2000US-0251990.
 PR 05-JAN-2001; 2001US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX Rosen CA, Barash SC, Ruben SM;
 WPI; 2001-483426/52.
 PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
 useful for preventing, diagnosing and/or treating cancers and
 metastasis -
 XX Disclosure; SEQ ID NO 31349; 3071pp + Sequence Listing; English.
 CC AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
 amino acid sequences given in AAK62170 to AAK91921. (i) have cytostatic
 activity, and can be used in gene therapy and vaccine production. (i)
 proteins and polynucleotides may be used in the prevention, diagnosis and
 treatment of diseases associated with inappropriate (i) expression. For
 example, they may be used to treat disorders associated with decreased
 expression by rectifying mutations or deletions in a patient's genome

PR 23-AUG-2000; 2000US-0227009.
 PR 30-AUG-2000; 2000US-0228924.
 PR 01-SEP-2000; 2000US-0229287.
 PR 01-SEP-2000; 2000US-0229343.
 PR 01-SEP-2000; 2000US-0229344.
 PR 01-SEP-2000; 2000US-0229345.
 PR 03-SEP-2000; 2000US-0229509.
 PR 03-SEP-2000; 2000US-0229513.
 PR 06-SEP-2000; 2000US-0230437.
 PR 06-SEP-2000; 2000US-0230438.
 PR 08-SEP-2000; 2000US-0231242.
 PR 08-SEP-2000; 2000US-0231243.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231413.
 PR 08-SEP-2000; 2000US-0231414.
 PR 08-SEP-2000; 2000US-0232080.
 PR 08-SEP-2000; 2000US-0232081.
 PR 13-SEP-2000; 2000US-0231968.
 PR 14-SEP-2000; 2000US-0232197.
 PR 14-SEP-2000; 2000US-0232398.
 PR 14-SEP-2000; 2000US-0232399.
 PR 14-SEP-2000; 2000US-0232400.
 PR 14-SEP-2000; 2000US-0232401.
 PR 14-SEP-2000; 2000US-0233063.
 PR 14-SEP-2000; 2000US-0233064.
 PR 14-SEP-2000; 2000US-0233065.
 PR 21-SEP-2000; 2000US-0234223.
 PR 21-SEP-2000; 2000US-0234274.
 PR 25-SEP-2000; 2000US-0234997.
 PR 25-SEP-2000; 2000US-0234998.
 PR 26-SEP-2000; 2000US-0235484.
 PR 27-SEP-2000; 2000US-0235834.
 PR 27-SEP-2000; 2000US-0235836.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.
 PR 29-SEP-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 02-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239935.
 PR 13-OCT-2000; 2000US-0239937.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 20-OCT-2000; 2000US-0241826.
 PR 01-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246529.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.

PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 06-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 11-DEC-2000; 2000US-0251990.
 PR 05-JAN-2001; 2001US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 XX Rosen CA, Barash SC, Ruben SM;
 WPI; 2001-465460/50.
 DR
 XX
 PT Novel polypeptides useful for diagnosing, treating, preventing and/or
 PT prognosing disorders related to the proteins, including cancers, immune
 XX disorders and neuronal disorders
 PS Claim 1; SEQ ID No 1391; 880pp; English.
 XX
 CC The invention relates to novel isolated polypeptides (I), and
 CC polynucleotides (II). (I), (II) and the antibody to (I) are useful for
 CC diagnosing, preventing and treating diseases including immune system
 CC disorders (e.g. congenital and acquired immunodeficiencies, autoimmune
 CC disorders (e.g. rheumatoid arthritis), inflammatory conditions, organ
 CC transplant rejections and graft versus host disease, infectious diseases
 CC (e.g. hepatitis C), bleeding disorders, haemoglobin abnormalities and
 CC other blood-related disorders (sickle cell anaemia), myeloproliferative
 CC disorders, primary haematopoietic disorders, hyperproliferative
 CC disorders (e.g. Gaucher's disease and cancer), neurodegenerative
 CC abnormalities (Down syndrome), ischaemic injury (e.g. stroke), renal
 CC disorders (e.g. glomerulonephritis), cardiovascular disorders
 CC (e.g. arrhythmia), respiratory disorders, dermatological disorders, in
 CC wound healing, epithelial cell proliferation, endocrine disorders (e.g.
 CC Addison's disease), reproductive system disorders, gastrointestinal
 CC disorder (inflammatory disorders), liver disorders (cirrhosis),
 CC as stimulators of B-cell responsiveness to pathogens, activators of
 CC T-cells, to induce higher affinity antibodies, and as a means to induce
 CC tumour proliferation in pathologies e.g. acquired immune deficiency
 CC syndrome (AIDS). AAS26976-AAS27850 represent novel signal transduction
 CC pathway protein coding sequences and PCR primers of the invention.
 XX
 SQ Sequence 11316 BP; 2843 A; 2769 C; 2919 G; 2785 T; 0 other;

Query Match 16.7%; Score 30; DB 22; Length 11316;
 Best Local Similarity 61.5%; Pred. No. 6.2;
 Matches .48; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

QY 42 AACCTCTTTTACACCGAATCGTCTGAGTCATCTTTGCCGCGAATATGGGACC 101
 |||| |||| |||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | |||

THIS PAGE BLANK (USPTO)